DOCUMENT RESUME

ED 205 569 TM 810 454

AUTHOR	McBryde, Bruce; Lamont, Ann
TITLE	Mapping of Assessment Procedures in Schools: A Study of Assessment Procedures in Three Subject Areas at Year 12 Level in Queensland State Secondary Schools.
REPORT NO	ISBN-0-7242-0795-3
PUB DATE	Oct 80
NOTE	146p.
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	English; *Evaluation Methods; Foreign Countries; Formative Evaluation; *Grade 12; Interviews; Secondary Education; Secondary School Mathematics; Second Languages; Student Attitudes; *Student Evaluation; Summative Evaluation; Teacher Attitudes; Teacher Made Tests
IDENTIFIERS	*Australia (Queensland)

ABSTRACT

In the light of recommendations for future directions in secondary education made by the Review of School Based Assessment in Queensland Secondary Schools, this study aimed to systematically map current internal assessment practices for grade 12, and to determine the degree of satisfaction in schools with these practices. The study was performed in two stages: (1) a preliminary pilot study of six schools to determine study strategy and procedures; and (2) a statewide survey involving twelve state secondary schools and three subject areas of English, mathematics and foreign languages. Data were collected by means of interviews with inspectors, principals, teachers and students. Documented within a section presenting the findings are descriptions of assessment objectives within schools; the types of assessment; assessment procedures and measures; evaluations of assessment practices and interviewee concerns and suggestions. It is concluded that there was satisfaction among administrators and teachers with current assessment procedures, but concern was evident that comparability between schools was not being achieved. Thus, further research into the validity of assessment procedures is recommended. Integrated within the study are thirty-six tables which illustrate findings. Extensive appendixes present interview schedules, summaries of responses and recommendations, and additional materials. (AEF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED205569

Mapping of Assessment Procedures in Schools

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
B McBryde

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A STUDY OF ASSESSMENT PROCEDURES IN THREE SUBJECT AREAS AT YEAR 12 LEVEL IN QUEENSLAND STATE SECONDARY SCHOOLS

By
Bruce McBryde
Ann Lamont

Research Branch,
Department of Education,
Queensland

October, 1980

TM810454

2/3

McBryde, Bruce
Mapping of assessment procedures in schools.

Bibliography.
ISBN 0 7242 0795 3.

1. Grading and marking (Students) — Queensland. I.
Lamont, Ann. II. Queensland. Dept. of Education.
Research Branch. III. Title.

373.12'64

TABLE OF CONTENTS

Page

LIST OF TABLES (ii)
LIST OF FIGURES (iv)
I Introduction 1
II Methodology 9
III Findings on School Assessment 20
  (1) Clarification of objectives within the school 20
  (2) Summative assessment 22
  (3) Formative assessment 35
  (4) Marking and validation procedures 44
  (5) Achievement and other exit measures 49
  (6) Attainment of the purposes of assessment 64
  (7) General concerns and suggestions 67
IV Summary 68
V Conclusions 80
Appendix 1: Interview Schedule - Inspectors 83
Appendix 2: Interview Schedule - Principals and Teachers 93
Appendix 3: Interview Schedule - Students 110
Appendix 4: Additional areas of concern and suggested changes 113
Appendix 5: Relationships between the recommendations of the select committee on education in Queensland, and the data collected during the Mapping of Assessment Procedures in Schools (MAPS) study 116
Appendix 6: Relationship between the recommendations of the Review of School-Based Assessment in Queensland Secondary Schools and the data collected in the Mapping of Assessment Procedures in Schools (MAPS) study 119
Appendix 7: Suggested further research 122
Appendix 8: The Board and its certificates 127
Appendix 9: Calculation of the Tertiary Entrance (TE) Score 131
Glossary of Terms 133
References 136

LIST OF TABLES

Table		Page
1	The Board's committees and their functions	2
2	Sampling frame showing: number of schools in each category; the theoretical sample; and the actual sample	14
3	The role of respondents in schools at the time of interview by their actual classification	16
4	Actual classification of the respondents in schools by the numbers in the whole population and in the sample	16
5	Level of decision making by percentages in the whole population and in the sample	17
6	Basis of teaching program by type of teacher	21
7	Basis of assessment program by type of teacher	21
8	Time (in minutes) devoted to each type of assessment per semester	24
9	Number of assignments given per semester by subject	24
10	Relative weight (per cent) given to various types of assessment	25
11	Types of items in the summative assessment programs by their use in schools and inspectors' recommendations for their use	27
12	Type of item by variation in percentage weight in the summative assessment programs	28
13	Type of item by variation in percentage of teacher-made items in the summative assessment programs	29
14	Percent of assessment that is common across classes by subject	30
15	Number of principals and teachers who wanted changes in the decisions relating to summative assessment	31
16	Number of principals and teachers who wanted changes in the decisions relating to weightings	32
17	Number of principals and teachers who wanted changes in the decisions relating to common/different tests	34

Table Page

18 Teacher perceptions of the delay in returning marked assessments by subject 37

19 Use of formative assessment by teachers 39

20 Desires of inspectors, principals and teachers regarding the amount of summative assessment 40

21 Students' desires regarding the number of summative assessments by subject 40

22 Variation in the number of summative assessments by subject 41

23 Mean testing time and assignment load by subject 41

24 Desires of inspectors, principals and teachers regarding the amount of formative assessment 42

25 Students' desires regarding the number of formative assessments by subject 43

26 Student satisfaction with comparability between classes by subject 45

27 Criteria used in determining cut-off marks on the M4 form 52

28 Perception of the nature of comments received at moderation meetings by type of respondent 53

29 Number of respondents satisfied with present moderation procedures 54

30 Teachers' perceptions of who prepared the Special Subject Assessment in their subject 58

31 Teachers' awareness of how the Special Subject Assessment is prepared in their subject 59

32 When standardisation occurred in combining semester scores to obtain the Special Subject Assessment for each subject 60

33 Respondents' perceptions of whether Special Subject Assessments can be manipulated 61

34 Respondents' perceptions of whether any groups of students are unfairly affected by the assigning of Special Subject Assessments 62

35 Number of schools using various descriptors on their end of semester school reports 63

36 Perceptions of principals and teachers of the present achievement of the purposes of assessment 66

LIST OF FIGURES

Figure		Page
1	The relationship of the Board to schools and school systems in Queensland	(v)
2	Sequence of events in the Board's moderation procedures	50

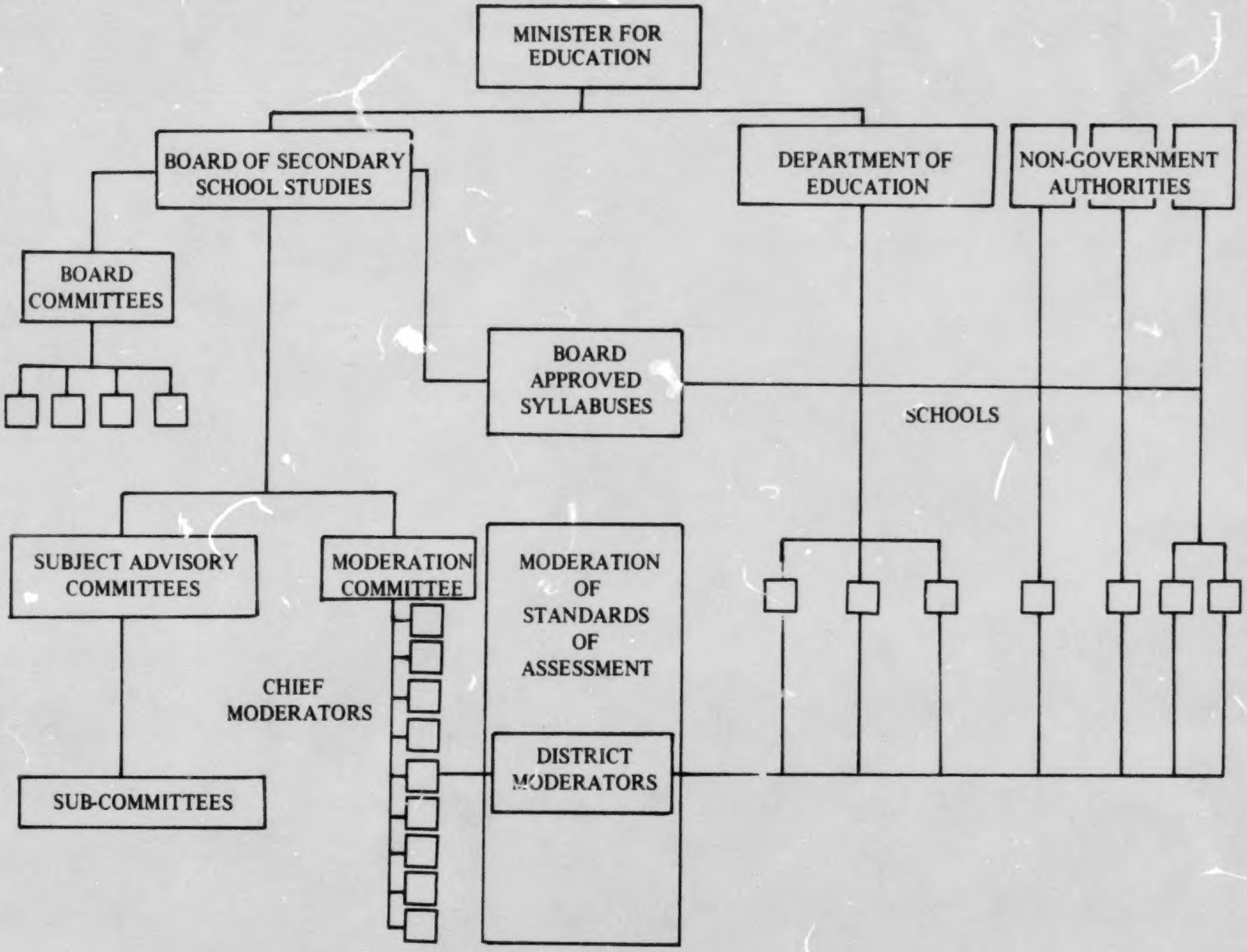
MINISTER FOR EDUCATION
BOARD OF SECONDARY SCHOOL STUDIES
DEPARTMENT OF EDUCATION
NON-GOVERNMENT AUTHORITIES
BOARD COMMITTEES
BOARD APPROVED SYLLABUSES
SCHOOLS
SUBJECT ADVISORY COMMITTEES
MODERATION COMMITTEE
MODERATION OF STANDARDS OF ASSESSMENT
CHIEF MODERATORS
DISTRICT MODERATORS
SUB-COMMITTEES

Figure 1: The relationship of the Board to schools and school systems in Queensland

I INTRODUCTION

BACKGROUND

Prior to 1971, performance of students in Queensland secondary schools was measured by their achievement in external examinations held at the end of Years 10 and 12. Entrance to tertiary institutions was determined by adding student ratings obtained by students in the Senior Public Examination held at the end of Year 12. In July 1969, the Minister for Education and Cultural Activities in Queensland established a committee chaired by Dr Wm C. Radford, then Director of the Australian Council for Educational Research, to review the system of examinations in Queensland and to make recommendations for the assessment of students' achievement. The following year this committee submitted its report *Public Examinations for Queensland Secondary School Students*, referred to as the Radford Report (Queensland Department of Education, 1970) to the Minister. The major recommendation of this report - to replace the two external examinations with a system of internal assessment - was accepted. This system was implemented in 1971 at the Year 10 level and carried through to Year 12 by 1973. The Board of Secondary School Studies was established to administer courses and certification. The relationship of the Board to schools and school systems in Queensland, together with the Board's committees and their functions have been described in information booklets distributed by the Board to schools (Information Kit, Unit 1, Board of Secondary School Studies, 1976). This description of the committees is reproduced in Table 1, and a diagrammatic representation of the relationship between the Board and schools is reproduced in Figure 1.

These changes were accompanied by broader curricula and a system of moderation meetings at which teachers from a district met to examine and discuss standards of work in order to achieve comparability between schools. In addition to these moderated assessments, schools were required to provide to the Board, a Special Subject Assessment (SSA) for each student in each Board subject studied. The assessments were assigned on a 1-99 scale and were not subjected to moderation. A State-wide order of merit list was prepared each year by rescaling these school assessments using results on the Australian Scholastic Aptitude Test (ASAT). Initially this was used to allot Commonwealth Scholarships but subsequently it became the method of determining tertiary entrance.

Table 1: The Board's committees and their functions

COMMITTEE	FUNCTIONS
SUBJECT ADVISORY COMMITTEES	To make recommendations to the Board concerning syllabuses devised by the committee or syllabuses devised elsewhere and presented to the Board for approval as Board subjects.
MODERATION COMMITTEE	To implement Board policies related to the moderation of standards of assessment for the award of Board certificates.
PUBLIC RELATIONS COMMITTEE	To advise the Board concerning the dissemination of information pertaining to the operations of the Board.
CURRICULUM COMMITTEE	To advise the Board on trends in secondary education syllabuses presented for Board approval.
ASSESSMENT & MODERATION COMMITTEE	To advise the Board on matters pertaining to assessment, moderation and the award of certificates.
MANAGEMENT & FINANCE COMMITTEE	To advise the Board on matters related to the management & financing of the operations of the Board.

The introduction of this system of internal assessment reflected a shift in philosophical emphasis within the education profession - the change being toward increased school responsibility rather than Department responsibility particularly for the curriculum, and assessment of students. This shift, however, required teachers to extend previously little used skills in order to manage the different procedures required with internal assessment. While this seems to have broadened the skills of most teachers, it has also made standards of work less visible to the community as a whole. Students in different schools no longer sit for a common examination covering the same content in each of a limited range of subjects. The number of subjects has increased with the introduction of school subjects as well as Board subjects. The content taught in the same subject differs from school to school, as do the assessment instruments used to measure achievement in that subject. The shift in responsibility for standards of achievement from an external examination board to individual schools has lead to some concern that overall standards may have changed and that standards may not be comparable from school to school.

While the system introduced in 1973 still operates, it has not remained static over this period. The requirements of the Board were rather sketchily conveyed at first. Gradually, a comprehensive set of forms was developed to meet all contingencies. In addition, there has been a shift in emphasis by tertiary institutions from the use of the moderated school assessment to the use of the Tertiary Entrance (TE) Score in determining entrance (see glosary of terms). This move towards the use of ASAT to rescale school assessment is not confined to Queensland. One other State and the A.C.T. have adopted procedures which use ASAT in the selection of tertiary entrants.

RELATED RESEARCH

Since the implementation of internal assessment, there has been much said and written by representatives of various groups in society by way of criticism and defence of that system. This had led to major research being undertaken by several groups into aspects of secondary schooling since 1971. This research includes a study by staff at the University of Queensland (Campbell et al, 1976), and the Schools Under Radford (SUR) Study, conducted jointly by the Department of Education and the Board of Secondary School Studies (Fairbairn, McBryde & Rigby, 1976). Both of these studies were based on data collected in 1974. In both cases, interviews and questionnaires were used. Questions ranged from pre-coded, objective items to those requiring open-ended responses. However, these were evaluation studies which have not investigated assessment procedures in detail. They were mainly concerned with a global evaluation of the system, one aspect of which is assessment. The main findings of these reports in the area of assessment are presented below.

Campbell et al (1976):

- students often regard ratings as the most important outcome of schooling;
- tests remain the dominant method of assessment although the use of other methods is increasing a little in many schools;
- there has been some liberalisation of the traditional framework in evaluation, but no radical new conceptualisation;
- such things as effort, attitude to oneself and others, and involvement, receive little acknowledgement in the assessment programs;
- tests are used to try to motivate students toward schoolwork;
- there have been increases in clarification of objectives and establishment of consistency among objectives, methods and assessment;
- teachers have experienced a greater sense of common involvement, more responsibility and autonomy, and, despite the strain, considerable growth in professional competence; and
- the Radford Scheme has generated stress and frustration, but satisfaction as well, and few would wish to revert to the system of public exams.

Fairbairn, McBryde & Rigby (1976):

- teachers did not believe that comparability was being achieved at moderation meetings;
- formative assessment was rarely used;
- there were long delays between when students submitted work and received marked scripts back;
- teachers felt they lacked expertise in test development;
- a high frequency of testing took place;
- while different tests between schools were considered acceptable, different tests between classes within a school were not;
- students were preoccuppied with ratings and relatively disinterested in locating their weaknesses;
- the majority of teachers wanted ratings to be issued less frequently while most students preferred ratings to be issued once per semester as was the case;
- principals' and teachers' workloads had increased;
- teachers were able to make their classroom assessment more relevant to the needs of their particular classes; and
- teachers were able to test a wider range of objectives.

In 1976, the Queensland Teachers' Union arranged a workshop to investigate moderation procedures in secondary schools in Queensland. The workshop and the subsequent publication concerning the workshop both bore the name *Moderation - at the crossroads?* (McMorrow [ed], 1976). Participants in this workshop came from a variety of educational institutions - the Queensland Teachers' Union, the University of Queensland, the Queensland Department of Education, the Schools Commission, the Board of Secondary School Studies and the Australian Council for Educational Research. Included were reviews of the findings of the Schools Under Radford Study and the report by Campbell et al. The overall observations of the reporting panel at the workshop with respect to assessment are summarised below.

McMorrow (ed), 1976:

> The difficulty of change in an established system was discussed. Hence, changes to assessment procedures should be accompanied by the active involvement of as many interested parties as possible, pre-service and in-service training courses designed to equip teachers for these changes, and an efficient communications network to monitor any changes.
>
> The last conclusion was that any change in assessment procedures should be preceded by full discussion of the options by as many different interested groups as possible.

In 1976, the Board of Secondary School Studies set up a Select Committee chaired by Professor E. Scott to review the report by Campbell et al and the Schools Under Radford Study and to advise the Board of their implications for Board policy and practices. This committee's report entitled *A Review of School-Based Assessment in Queensland Secondary Schools* contained recommendations as to possible future directions for secondary education (Board of Secondary School Studies, 1978). The major recommendations relating to assessment are given below.

Assessment of Student Achievement:

P15 The present norm-based assessment procedures in Years 10, 11 and 12 should be replaced by competency-based procedures

P19 In the Senior Secondary School, the Board should issue only one assessment in each subject - an 'exit assessment' indicating a student's global level of performance in that subject at the cessation of his/her study of that subject.

Accreditation and Certification:

P20 The Board's present moderation practices should be replaced by a system of 'subject accreditation' of proposed work programs and competency criteria, and 'certification' of reported standards of achievement.

P24 The present moderation meetings in Years 11 and 12 should be replaced by 'district teachers meetings' in each subject or subject grouping. Such meetings should be held twice a year during the mid-semester breaks in the first and second semesters respectively. Attendance at these meetings by 'school subject co-ordinators' (replacing the present position of school subject moderators) should be mandatory. However, such meetings should not be seen as part of the formal accreditation and certification processes, but rather as professional meetings for teachers to engage in co-operative ventures in planning their work and assessment programs and in arriving at performance standards seen as fulfilling competency criteria.

Maintenance of Standards:

P26 To assist in the maintenance of State-wide achievement standards and the maintenance of such standards across time, the spirit of paragraph 6.15 of the Radford Report should be endorsed. A policy should be adopted by using 'competency reference tests' in Board subjects. The sole objective of these tests should be to assist schools in determining standards of performance relative to each level of competency in a subject. Such tests will be an invaluable aid to teachers in determining competency standards in smaller schools. Under no circumstances should the results of competency reference tests appear on Board or school certificates or reports. Initially, competency reference tests might be restricted to the Senior Secondary School with extension downwards to Year 10 depending upon the subsequent advice of the Accreditation and Certification Committee.

Tertiary Entrance Score:

P27 For the purpose of determining order of merit for entry to tertiary institutions, the Tertiary Entrance Score should be retained, though the Board should continue its research into the efficacy of the Tertiary Entrance Score as a method of ranking students.

P28 In calculating the Tertiary Entrance Score use should be made of ASAT, or of a comparable test, as at present. However, the Board should continue its research into the use and efficacy of such a moderating instrument.

Certificates:

P30 Formal reporting to students and parents and formal certification of student performance should be the responsibility of both the school and the Board of Secondary School Studies. The Board should be responsible for issuing a junior certificate and a senior school certificate at the end of Years 10 and 12 respectively. Such certificates should report only in terms of academic achievement. The school should provide additional information in those areas of performance and development which are not the Board's concern, viz. leadership, sporting ability, attitudes, values, social developments, emotional maturity. Thus, the school should be encouraged to issue a school leaving certificate containing such additional information to each student at the time of his/her leaving school.

The continuing concern with education among various groups within the community led to a Select Committee on Education in Queensland being appointed in April 1978 by the Queensland Parliament to inquire into, report upon and make recommendations in relation to the system of education in Queensland and the extent to which it meets the expectations of students, parents and the community. Their First Interim Report was completed in 1978 (Queensland Legislative Assembly, 1978).

AIMS OF THE PRESENT STUDY

At the time this study was initiated, the researchers felt that there were a number of important areas in which research could be profitably undertaken. These included assessment, curriculum issues, teaching techniques and school administration. Although assessment was chosen, it cannot be inferred that it was considered the most significant aspect of the educational process. However, decisions about the nature of exit assessments are made by bodies outside the school and are imposed on the school. These decisions affect the options available to schools in their decision making about other aspects of their operation. Thus, although assessment may be only one aspect in the total educational process, the decisions in this area have a major impact on all other aspects of the school's program.

Although both the Schools Under Radford Study and the report by Campbell et al indicated that there was considerable dissatisfaction in schools in the area of assessment, neither focussed specifically on assessment, nor on detailed practices within particular subjects. In view of the changes to assessment procedure being presently considered by the Queensland Government as a result of the review of school based assessment, it appeared desirable to obtain a more specific picture of current assessment procedures in State secondary schools. As well as providing useful information at this point in time, this data may be used for comparative purposes if substantial changes occur in assessment procedures.

The researchers sought to provide two kinds of information - firstly, factual information on major assessment procedures and issues, and the decision making associated with them, and secondly, evaluative comment on each of these. In this study, the researchers have attempted to focus on some specific issues within the general area of assessment. To achieve this, semester four at the Year 12 level was chosen and only three subject areas were investigated.

The major aims of the study were as follows:

- (i) to map the decisions relating to the assessment procedures used at the Year 12 level within State secondary schools;
- (ii) to determine the degree of satisfaction with present procedures amongst those involved with assessment in schools; and
- (iii) to examine the effects of the requirements of outside bodies (the Department of Education, the Board, tertiary institutions and employers) on the assessment programs within the school.

Changes in assessment practices and the satisfaction of those involved during the period of time that internal assessment has been used in Queensland have not previously been examined. In addition, there has been criticism that the present assessment practices do not meet the purposes of assessment. Thus, it was considered desirable that comparisons be made and that the views of respondents on the attainment of the purposes of assessment be obtained.

Additional aims were:

(a) to compare the assessment situation in 1978 with the corresponding situation in 1974 as far as is possible with the data available; and

(b) to determine the extent to which school staff believe that the present system facilitates the achievement of the purpose of assessment in these schools.

II METHODOLOGY

The study was conducted in two stages - the first, a preliminary stage involving six schools at various levels of participation; the second, a State-wide survey involving 12 State secondary schools.

STAGE 1 - THE PILOT STUDY

The aims of Stage 1 were to identify the assessment issues to be investigated in Stage 2, to decide on strategies and techniques to be used, and to test the instruments developed for Stage 2 of the study. For this stage of the study, five metropolitan high schools and one country high school were used.

The initial list of issues to be explored was drawn up on the basis of initial perceptions of relevant educational issues on the part of the researchers.

In the first two trial schools, open-ended discussions with the principal and subject masters were used to modify this list of assessment issues to include issues of concern to school staff. The resultant list is given below:

- clarification of objectives within the school
- summative assessment
- formative assessment
- marking and validation procedures
- achievement and other exit measures
- attainment of the purposes of assessment
- general concerns and suggestions

The approach selected by the researchers was to trace decisions and actions from the time when teachers clarified their objectives in the development of their programs of work and assessment to the production of the output measures of student characteristics (the moderated school ratings, the assigning of the SSA for the calculation of the TE Score, and the school report). The principals, subject area co-ordinators and other teachers make different contributions to the decisions and actions relating to the school's assessment program. It was expected therefore that both the involvement and the degree of satisfaction could vary across these groups, and for this reason they were identified in the design of the survey.

In the next two trial schools, structured interviews were held separately with the principal, subject masters and other teachers in three subject areas. These interviews were used to determine the details of each of the issues to be investigated and to formulate the questions to be asked of principals and teachers.

The detailed list of issues selected for investigation is as follows.

Clarification of objectives within the school:

- relating to the whole school
- relating to specific subjects within the school

Summative assessment:

- when assessments are administered
- time allowed
- weighting on various assessment elements
- weighting on types of items
- weighting on modes of assessment
- use of common tests
- setting of assessment instruments

Formative assessment:

- amount of formative assessment
- use of summative assessment for formative purposes
- balance between summative and formative assessment

Marking and validation procedures

Achievement and other exit measures:

- moderated school assessments
- special subject assessments
- school reports

Attainment of the purposes of assessment

General concerns and suggestions

The following aspects of each of the first five issues were examined.

- What decisions are implemented?
- Who makes the decisions?
- Why are decisions made?
- Are decisions satisfactorily communicated?
- What are the perceived constraints?
- What would the respondents prefer?
- What are the advantages and problems associated with preferred alternatives?

The specific questions formulated to probe these areas form the principals' and teachers' interview schedule which is included in this report as Appendix 1.

From discussions in the first four trial schools, it became evident that the views of students on relevant issues would make a worthwhile contribution to the study. The remaining two schools were used to provide the students' interview schedule. Discussions with groups of students combined with the judgement of the researchers indicated that the relevant issues were:

- amount of formative assessment;
- use of summative assessment for formative purposes;
- balance between summative and formative assessment; and
- marking procedures.

Again, the specific questions formulated on these issues constituted the interview schedule, and this is contained in Appendix 2.

It was decided as a result of Stage 1, that it would be desirable to interview inspectors involved in the subject areas selected. It was felt that because of their experience both with the subject area itself and as school administrators, inspectors would be able to provide a broad perspective on the major issues raised with school staff, although the aspects of these issues which were specific to particular assessment programs were not applicable. The relevant questions from the principals' and teachers' interview schedule formed the inspectors' interview schedule. This is presented in Appendix 3.

To summarise, the same interview schedule was administered to various levels of staff within the school, i.e. the principal, subject master and teachers. The interview schedule for inspectors was a subset of this. The student interview schedule examined some of the issues covered with principals and teachers.

All three interview schedules were semi-structured. They contained both specific questions requiring objective answers which were often precoded on the schedule, together with open ended questions which gave respondents an opportunity to elaborate on issues which were significant to them. At the end of each interview, time was allowed for respondents to raise any issue of particular importance which had not been covered during the interview.

STAGE 2 - THE MAIN STUDY

Sample

Due to time constraints in the collection of the data it was decided to limit the study to a few subject areas in a small sample of schools. It was felt that an indepth study of three subject areas in 12 schools would provide a sufficiently comprehensive picture of assessment procedures in these subjects and the effect on these subjects of decisions which apply to the whole school.

Subjects

The choice of subject areas to be investigated in the study was not easy. Each Board subject has its own unique features which would justify its inclusion in the study. English and Maths I were chosen because they were taken by large numbers of students and the skills developed in these subjects differ from each other and were used by students in a wide range of other subjects. Foreign languages was chosen because each foreign language is a subject studied by a small group which is purported to be drawn from a relatively elite group of high achievers. A subject area containing such groups was chosen because it is claimed that small groups face different problems from large groups and that ratings are more frequently disputed when groups of high achievers are involved.

Since the subjects were chosen according to these specific criteria, the results of this study will provide a valid picture of assessment in these three subjects only, rather than one that can be generalised across all subjects.

Schools

It was believed that information from a representative sample of 12 schools would provide a reliable picture of assessment patterns within the chosen subjects across the State. In addition, this was the largest sample which could be visited in the time available in third term in 1978, allowing time for an interim report to be written before the end of the school year.

The sampling of schools was a two-way stratified classification based on the size of the school and its geographic location. Schools were classified as small, medium or large, based on their enrolment at January 1978 using cut-offs of 525 students and 1,000 students. Schools were also classified as metropolitan, provincial city or country. Schools within the Brisbane urban area were classified as metropolitan. Where a city was large enough to have two or more State high schools, these were classified as provincial city schools. The remaining schools, i.e. those in towns with only one State high school or secondary department were classified as country schools.

As it was possible that school size would be a factor in the degree of involvement of respondents in decision making and in the ease of communication within the school, a representative sample with respect to size was considered desirable.

The moderation process requires communication between schools and with the Board. Since communication between schools and communication with the Board and the Department of Education is likely to be affected by location, stratification with respect to location was considered desirable. In the metropolitan area, other schools were nearby as were the Board and the Department of Education. In provincial cities, other schools were nearby but the Board and the Department of Education were not, while in country towns, other schools were widely spread as well as the Board and the Department of Education being distant.

State high schools without Year 12 students were not included in the classification of schools, while State primary schools with secondary departments having Year 12 students were included. The number in each classification is shown in Table 2.

The next step in the sampling procedure was to calculate the number of sample schools to be chosen from each cell. Twelve State schools were being selected from a total of 134 (i.e. a 9.0 per cent sample). This proportion of the number of schools in each category was calculated. These values were rounded to whole numbers to obtain the actual number of schools to be selected from each cell. The theoretical and actual sample sizes for each cell are also shown in Table 2.

Table 2: Sampling frame showing: number of schools in each category; the theoretical sample; and the actual sample

Location	Size			TOTALS
	Large (> 1000 students)	Medium	Small (< 525 students)	
Metropolitan				
Number of schools	20	17	2	39
Theoretical sample	2.0	1.7	0.2	3.8
Actual sample	2	2	0	4
Provincial City				
Number of schools	18	7	1	26
Theoretical sample	1.8	0.7	0.1	2.6
Actual sample	2	1	0	3
Country				
Number of schools	3	16	38	57
Theoretical sample	0.3	1.6	3.7	5.6
Actual sample	0	1	4	5
Totals				
Number of schools	41	40	41	122
Theoretical sample	4.0	3.9	4.0	12.0
Actual sample	4	4	4	12

To obtain the actual schools to be visited, schools were grouped into clusters which were within easy driving distance from each other. Clusters were then selected to provide a representative cross-section of geographic locations while minimising the costs of travel. The areas selected were North Queensland, Darling Downs, Wide Bay and the metropolitan area. Schools were then chosen at random within these specifications. This sample was not representative with respect to educational region (and thus not strictly representative of geographical location), but it was representative of the *types* of geographical region referred to above.

Respondents

The research design required that the principal, the subject area co-ordinator (SAC) and another teacher at Year 12 level in each of the three subject areas studied, and a group of Year 12 students be interviewed in each school visited. However, in all schools, at least one of the categories was not appropriate. No school has two classes studying the same foreign language, thus, the other teacher in this category was not relevant. In some schools, there was only one teacher in either or both of English and Maths I. In the 12 schools collectively, the following categories were relevant:

Principals	12
English SAC	12
English 'other teacher'	9
Maths I SAC	12
Maths I 'other teacher'	8
Foreign language SAC	11
TOTAL	64

In many cases, the people performing these roles had been appointed to positions in the school in an acting capacity. e.g. of the 12 'principals', five were deputy principals who were acting as principal during the prolonged absence of the principal.

In two cases, one respondent filled two of the 64 roles - one interviewee was the subject area co-ordinator in both English and French, and another interviewee was both the principal and Maths I subject area co-ordinator. These respondents were interviewed in both capacities. There were several interviews where two respondents gave data jointly for a single role (e.g. two Year 12 English teachers interviewed together). Consequently, 64 interviews with school staff were conducted although the actual number of participants in these interviews did vary from 64. Table 3 shows the actual classification of the respondents involved in each of the 64 interviews.

The teaching experience of the teachers is shown in Table 4. Where several teachers contributed to the one interview, the classification and experience of the person making the major contribution to the interview were recorded. It should be noted that the respondents were drawn mainly from experienced teachers within the schools. In only 11 of the 64 interviews in schools did the major respondent have fewer than six years experience. This reflects the extent to which administrators and teachers responsible for Year 12 students are drawn from experienced teachers. As indicated in Table 5, the proportions of subject masters and school administrators interviewed greatly exceed the corresponding proportions for the State. However, it was not intended that the sample be representative of experience and teaching classification, but rather that it should reflect levels of decision making affecting assessment at Year 12.

Table 3: The role of respondents in schools at the time of interview by their actual classification

		Principal	Deputy Principal	Senior Mistress	Subject Master	Teacher	TOTALS
Principal		7	5	0	0	0	12
SAC *	English	0	0	0	9	3	12
	Maths I	1	0	0	8	3	12
	Foreign Languages	0	0	0	1	10	11
'Other Teacher'	English	0	1	1	0	7	9
	Maths I	0	0	0	0	8	8
TOTALS		8	6	1	18	31	64

* Subject Area Co-ordinator

Table 4: Actual classification of the respondents in schools by the numbers in the whole population and in the sample

ACTUAL CLASSIFICATION		POPULATION (State-wide)		SAMPLE (MAPS sample)	
		Number	Per cent	Number	Per cent
Teacher	1st year	594	8	0	0
	2nd year	682	9	3	5
	3rd to 5th year	1953	26	8	13
	6th or later year	3109	42	20	31
Subject master		803	11	18	28
Senior mistress		116	2	1	2
Deputy principal		139	2	6	9
Principal		122	2	8	13
TOTALS		7518	100 *	64	100 *

* These columns do NOT actually add to 100 because of the rounding of individual percentages.

Table 5: Level of decision making by percentages in the whole population and in the sample

LEVEL OF DECISION MAKING	POPULATION (Per cent)	SAMPLE (Per cent)
Teacher	84	48
Subject master	11	28
Administrator	5	23
TOTALS	100	100 *

* This column does NOT actually add to 100 because of the rounding of the individual percentages.

In addition to the interviews with school staff, a class of Year 12 students was interviewed in each school, the interview schedule for students probed the areas from the principal's and teacher's interview schedule which were relevant to students. This enabled the researchers to determine the extent to which the teachers' perceptions matched those of the students in relation to assessment in the three specific subject areas within their school.

Schools were asked to include as many Maths I and foreign language students as possible in the student group to be interviewed. (All students studied English.) They were asked to select a student group which satisfied these criteria while minimally disrupting the school. Thus, the students in the group selected were not necessarily taught those subjects by the teachers interviewed. A total of 222 students were involved in these interviews.

Most of the issues raised with inspectors related to assessment in their subject area. It was assumed that their experience as teachers and administrators in schools together with their experience in visiting a large number of schools in a professional capacity would allow them to base their opinions on the global view of the problems and issues. Six inspectors were interviewed - three being involved with mathematics, two with English, and a further one with both English and foreign languages.

Of the six inspectors interviewed, four were members of subject advisory committees or other Board committees involved in syllabus implementation or development. Thus, the inspectors' schedule was designed to obtain authoritative opinions from a different perspective from that of the school staff. Overall, three inspectors were interviewed in the area of English, three in mathematics and one in foreign languages (one was responsible for both English and foreign languages).

Interviews

As a result of the data collected during Stage 1, it was decided to use interviews rather than questionnaires. It was felt that significant nuances relating to the issues would be lost in the questionnaire approach.

It was estimated that from 60 to 90 minutes would be required for each principal and teacher interview, and about 30 minutes for the student interview. The number of interviews required varied from seven in large schools to four in small schools. Thus, the researchers spent one day in each of the four small schools and two days in each of the medium and large schools.

The program of interviews was organised by the schools. As previously noted, the researchers specified the criteria for selecting interviewees, but the actual selection and timetabling was left to individual schools. This was done to minimise the disruption caused in the schools. This policy did give schools freedom in the choice of teachers other than the subject area co-ordinators and in the Year 12 class interviewed. However, it is possible that a systematic bias may have been introduced by the method of selection of these categories.

During all interviews, both researchers were present. The technique used required that each researcher should be responsible for clarifying the issues and making notes before moving on to a new topic. Although the duplication of note-making may have been unnecessary for some questions, it proved to be valuable for others. This use of two interviewers assisted the researchers in clarifying misconceptions and ambiguous statements during the interview before the data were recorded. The researchers believe in addition, that the process was invaluable in ensuring that ambiguities were avoided and that the validity of the data would be high. The use of two interviewers also facilitated post-interview discussions which were used in discussing the results obtained, and in clarifying the nature of conflicting data obtained during different interviews.

In the student interviews, and those teacher interviews where two respondents were present, it was made clear that respondents should answer for themselves, and not just conform to a majority viewpoint. Where there was not agreement in these interviews, the number of respondents expressing each view was recorded.

While satisfied with the techniques used in the study, the researchers did find that, in some schools the time available was too short to pursue certain points at even greater depth.

Data Collation

As mentioned previously, the interview schedules required both objective answers which were precoded, and open ended answers. The responses of each type of interviewee (inspectors, principals, subjects area co-ordinators in English, Maths I and foreign languages, 'other teachers' in English and Maths I, and students) were collated separately. On the precoded objective questions, the responses were tallied producing a total for each possible option.

The researchers discussed the open ended responses and subsequently summarised, coded and grouped them. These too, were tallied where appropriate. Minority and conflicting viewpoints were recorded so that a complete picture of respondents' attitudes could be obtained.

Broad comparisons are made between assessment practices in 1974 and 1978. However, no statistical tests can be applied in this respect. This is due to the subjective nature of the data and the lack of consistency in foci of the studies under consideration.

III FINDINGS ON SCHOOL ASSESSMENT

(1) CLARIFICATION OF OBJECTIVES WITHIN THE SCHOOL

Bloom, Hastings & Madaus (1971, p 21) comment that *once objectives are defined clearly, they can become models or plans that help shape and guide the instruction and evaluation processes.* They perceive that the clarification of objectives is an essential link between teaching and evaluation.

The change to internal assessment has caused this clarification of objectives to become more important within the school, both at the broad school level and at the individual subject level.

School policies related to assessment

All schools had definite policies relating to some aspects of assessment. This usually took the form of maximum limits on the frequency and/or length of assessments.

In some schools, binding assessment procedures were prescribed by the school administration in the form of a handbook provided to all staff. In most other schools, respondents followed a system which was regarded as the normal practice in their school. In a previous study in which over 80 per cent of principals in the State responded (Fairbairn & Johnston, 1978) principals were asked whether their school-wide assessment procedures were 'binding', 'normal practice' or 'neither binding nor normal practice'. The vast majority of principals (73 per cent) indicated that a normal practice existed in their schools, while the remainder was equally divided between the first and last options.

Course Objectives

All schools are required to submit to the Board an M1 form (see glossary of terms) for each subject. Included on this form is a statement of the assessment objectives for the course of study at that school. Often a more detailed statement of course assessment objectives occurs in other school documents such as the workbook.

In all schools, the information on the M1 form was prepared by the subject area co-ordinator. In some schools, other teachers assisted with its preparation. No teacher expressed concern about not being involved with its preparation.

All teachers except one indicated that the M1 form was accessible to them when they needed to refer to it. Most teachers said that they referred to their M1 forms between two and six times per semester. However, the teachers pointed out that the contents of the M1 form were contained in their workbooks to which they referred more frequently.

The majority of teachers teach from their workbook, although some of them use other bases as well. 'Other teachers' in Maths I tend to use textbooks as a basis for their teaching (see Table 6).

Table 6: Basis of teaching program by type of teacher*

BASIS OF TEACHING PROGRAM	ENGLISH		MATHS I		FOREIGN LANGUAGE SAC (11)
	SAC ** (12)	'Other teacher' (9)	SAC (12)	'Other teacher' (8)	
Syllabus	2	1	3	0	2
M1 form	2	1	2	2	3
Workbook	8	7	9	3	6
Textbook	1	0	3	7	4
Other	2	2	0	1	2

* One or more responses were allowed. Therefore, the total of each column may not equal the number of schools in which each type of teacher was interviewed which is shown in brackets.

** Subject Area Co-ordinator

Table 7 shows that the M1 form is usually used as the basis of the assessment program. However, the workbook in English and the teaching program in Maths I and foreign languages are also used extensively.

Table 7: Basis of assessment program by type of teacher*

BASIS OF ASSESSMENT PROGRAM	ENGLISH		MATHS I		FOREIGN LANGUAGE SAC (11)
	SAC (12)	'Other teacher' (9)	SAC (12)	'Other teacher' (8)	
Teaching program	2	1	3	3	4
Syllabus	1	0	1	0	3
M1 form	8	6	7	2	4
Workbook	5	5	4	1	1
Textbook	0	0	0	1	1
Other	0	0	1	0	0

* Refer to comments on Table 6

Inspectors commented that the workbook is the school's interpretation of the syllabus and thus the two should be closely linked. They also felt that the workbook and the teaching program should be equivalent and that the assessment program should be fully detailed in the workbook.

(2) SUMMATIVE ASSESSMENT

Summative assessment includes all tests, assignments, practical and other work which contribute to the marks used in measuring or rating the student's achievement.

The following aspects of summative assessment are discussed in this section.

- When tests are administered
- Time allowed
- Weighting on various assessment elements
- Weighting on modes of assessment
- Weighting on types of items
- Use of common tests
- Setting of assessment instruments

The chapter is divided into two parts - firstly, the structure of the summative assessment program is discussed with respect to the above aspects; and secondly, the decision making processes involved are discussed using similar headings.

Structure of the summative assessment programs

For each subject in each school, teachers were asked to outline the extent and nature of their summative assessment programs. Some patterns emerged showing both similarities and differences among subjects. Inspectors were asked for their views on what constituted an appropriate summative assessment program.

When tests are administered

In Maths I, the ratings are based almost entirely on formal tests under examination conditions. All schools visited had an end of semester test with a weighting of at least 40 per cent. All but one of these schools also had a mid-semester test which accounted for, from 20 to 40 per cent weighting. In most schools, there were no other elements in the summative assessment program in Maths I. The major exception was the computer maths unit where assignments were part of the assessment program. The attitudes of inspectors on this issue corresponded closely with the present practice.

In English, the ratings are based on a large number of assessments of a widely varying nature. All schools visited had at least one test under examination conditions at the end of the semester, and at least one such test during the semester. In all schools, there was an oral/aural component and a component for unsupervised work completed at home. Generally, assessment in English occurred more frequently during the semester than in Maths I.

In foreign languages, the ratings are also based on a large number of assessments of a widely varying nature. The assessment structure is similar to that for English with a greater emphasis on oral/aural assessment. In general, assessment occurs more frequently in the foreign languages than in either English or Maths I.

Teachers and inspectors were asked whether there were any course objectives which could not be adequately assessed. The affirmative responses of both groups fell into two categories, viz. the general objectives relating to appreciations and attitudes, and those objectives which are currently being assessed with difficulty, e.g. oral/aural work.

The frequency of assessments suggested by inspectors in English and foreign languages broadly corresponded with the present assessment practices in these subjects.

Time allowed

The time devoted to assessment varied considerably from school to school. This is shown in Table 8. The total time per semester devoted to formal tests written under examination conditions was similar for the three subject areas covered, although it was slightly higher for English (average, 252 minutes) than for foreign languages (average, 210 minutes) and Maths I (average, 202 minutes).

However, for Maths I, this represented the total summative assessment program, but students of English had, in addition, from one to eight assignments to be prepared at home, and up to two assignments to prepare in class time (see Table 9). In addition, oral/aural assessment in English occurs informally during the semester. The burden of extra assessment is not as great in the foreign languages where an average of one extra assignment is required, together with about 55 minutes of oral/aural assessment. Thus, it is not surprising that the small number of students who are not satisfied with the present assessment load tend to want a decrease in English and an increase in Maths I. The issue of student satisfaction with the assessment load is expanded in the third section of this chapter.

Table 8: Time (in minutes) devoted to each type of assessment per semester

		Minimum	Maximum	Median	Mean
(a) ENGLISH					
Written under examination conditions	DURING	60	200	105	108
	END	80	240	165	144
ORAL/AURAL		Mostly assessed informally through the semester.			
TOTAL		170 *	390 *	277 *	252
(b) MATHS I					
Written under examination conditions	DURING	50	135	60	70
	END	90	180	150	132
ORAL/AURAL		0	0	0	0
TOTAL		150 *	285 *	205 *	202
(c) FOREIGN LANGUAGES					
Written under examination conditions	DURING	0	210	75	91
	END	90	205	120	119
ORAL/AURAL		30	90	53	55
TOTAL		165 *	410 *	255 *	265

* The minimum total testing time cannot be found by simply adding the individual minima as these may not all occur within the same school. This also applies to the maximum total testing time and the median total testing time.

Table 9: Number of assignments given per semester by subject

		Minimum	Maximum	Median	Mean
ENGLISH	IN CLASS	0	2	0	1
	AT HOME	1	8	3	3
MATHS I	IN CLASS	0	2	0	0
	AT HOME	0	2	0	0
FOREIGN LANGUAGES	IN CLASS	0	2	0	0
	AT HOME	0	5	0	1

Weighting on modes of assessment

The average weighting given to written examinations varied from 66 per cent in English and foreign languages to 100 per cent in Maths I (see Table 10). The only exception to the use of written examinations in Maths I appears to be in the unit on computing. Some mathematics teachers commented that they used assignments in social mathematics, although this subject was not included in the present study. (However, one inspector recommended the use of assignments in other sections of mathematics as well, while another mathematics inspector thought assignments should be used for formative but not summative assessment.) The other modes of assessment used were mainly assignments prepared at home in English (10 to 43 per cent), and oral/aural in the foreign languages (15 to 35 per cent). (One English inspector cautioned against the use of 'at home' assignments.)

Table 10: Relative weight (per cent) given to various types of assessment

		Minimum	Maximum	Median	Mean
(a) ENGLISH					
Written under examination conditions	DURING	10	50	23	27
	END	29	60	40	39
ORAL/AURAL		3	10	7	7
Work supervised IN CLASS		0	20	0	4
Work unsupervised AT HOME		10	43	20	23
(b) MATHS I					
Written under examination conditions	DURING	0	50	34	33
	END	50	100	66	67
ORAL/AURAL		0	0	0	0
Work supervised IN CLASS		0	3	0	0
Work unsupervised AT HOME		0	3	0	0
(c) FOREIGN LANGUAGES					
Written under examination conditions	DURING	0	45	25	23
	END	25	70	40	44
ORAL/AURAL		15	35	30	29
Work supervised IN CLASS		0	10	0	1
Work unsupervised AT HOME		0	15	0	3

Data collected in 1974 for the Schools Under Radford Study suggested that teachers at that time were not taking advantage of the flexibility in assessment procedures afforded by the implementation of the Radford Report.

> *Two reasons were given by teachers to explain the limited use of assessment techniques which would cater for individual differences:*
>
> - *moderation meetings won't accept assignments or oral work because they are too subjective; and*
> - *schools have flexibility in the use of assessment techniques but they are not taking advantage of it.*
>
> (Schools Under Radford Study, p 20)

In 1978, some teachers seemed to be taking more advantage of this flexibility than they did in 1974. However, although moderation meetings accept the use of some non-written assessments by schools when proposing students' ratings, they still appear to base their judgements on written work alone when determining the distribution of ratings for the school. English and foreign language teachers appear to be basing their assessment on those proposed in paragraph 5.16 in the Radford Report.

> *The schools, within the broad framework of an approved syllabus, and in accordance with clearly stated objectives for the syllabus, will have freedom to choose detailed content, freedom in method of presentation and, within the moderation system, freedom in methods of assessment of student performance.*
>
> (Radford Report, p 70)

In these subjects the teachers would appear to have capitalised on the flexibility available through internal assessment.

Weighting on types of items

The types of items used and the weightings given to these varied among subjects. This is shown in Tables 11 and 12. In all schools, extended response items (e.g. essays or maths problems) were used in all three subjects (see Table 11).

Table 11: Types of items in the summative assessment programs by their use in schools and inspectors' recommendations for their use

Type of Item		(a) ENGLISH		(b) MATHS I		(c) FOREIGN LANGUAGES	
		Frequency of Use in Schools (12) *	Frequency of Inspectors' Recommendations for Use (3) +	Frequency of Use in Schools (13) **	Frequency of Inspectors' Recommendations for Use (3)	Frequency of Use in Schools (11) ***	Frequency of Inspectors' Recommendations for Use (1)
Multiple choice		8	2	2	0	8	1
Completion (short answer)		8	2	6	3	10	1
True/false		1	0	0	0	2	0
Extended response (Essay/maths problem)		12	3	13	3	11	1
Assignment	In class	9	3	1	1	2	1
	At home	12	3	1	1	2	1
	Group	3	3	0	1	1	1
Practical work	Debate	8	3	0	0	0	0
	Drama	8	3	0	0	0	1
	Conversation	8	3	0	0	9	1
	Listening Comprehension	4	3	0	0	11	1
	Dictation	0	1	0	0	7	1
	Reading a Passage	3	1	0	0	7	1
	Lecturette	3	0	0	0	3	0
	Other	2	0	0	0	0	0

* Numbers in brackets indicate the number of school assessment programs examined.

** In one school, data was collected on two semester units which were being taught concurrently.

*** One school had no foreign language students in Year 12.

+ Numbers in brackets indicate the number of inspectors interviewed in that subject.

In Maths I, these were used almost exclusively. In six schools, a small proportion of the assessment in Maths I was made up of short answer questions, and two schools used some multiple choice items in Maths I.

In English and foreign languages, a wide variety of types of items was used in most schools. In these subjects at least, it appears that teachers are taking advantage of the flexibility available under the present system of internal assessment. The only areas in which inspectors' recommendations appeared to differ substantially from the practice in schools, were group assignments in English and listening comprehension in English. In both cases, the three English inspectors recommended their use but relatively few teachers were using them.

Multiple choice items were used in English and foreign languages in the majority of schools; however, these constituted a very small percentage of the overall assessment program (see Table 12). Completion or short answer items were used to a small extent as well in these two subjects, while true/false items were seldom used. The average weighting on assignments was 26 per cent in English but only five per cent in foreign languages, while the average weighting on oral/aural in English was small (seven per cent) but a significant proportion (27 per cent) in foreign languages. This latter result could be a reflection of the differing emphases placed on oral/aural assessment by the syllabus and moderation meeting between English and foreign languages.

Table 12: Type of item by variation in percentage weight in the summative assessment programs

Type of Item		Minimum	Maximum	Median	Mean
(a) ENGLISH					
Multiple-choice		0	17	5	6
Completion (short answer)		0	30	10	9
True/false		0	3	0	0
Extended response (Essay/maths problem)		30	84	47	53
Assignment	In class	0	20	6	8
	At home	6	43	16	18
	Group	0	0	0	0
Practical work/oral		3	10	6	7
(b) MATHS I					
Multiple-choice		0	4	0	1
Completion (short answer)		0	21	0	6
True/false		0	0	0	0
Extended response (Essay/maths problem)		75	100	100	94
Assignment	In class	0	0	0	0
	At home	0	6	0	1
	Group	0	0	0	0
Practical work/oral		0	0	0	0
(c) FOREIGN LANGUAGES					
Multiple-choice		0	20	5	6
Completion (short answer)		0	40	20	18
True/false		0	11	0	2
Extended response (Essay/maths problem)		20	70	40	40
Assignment	In class	0	10	0	2
	At home	0	15	0	3
	Group	0	0	0	0
Practical work/oral		0	30	30	27

Origin of test items

Teachers were asked whether they prepared their own items or whether they obtained them from other sources. In recording the data, it was sometimes difficult to classify responses, as many teachers, particularly in English and foreign languages indicated that they obtained their ideas from other sources, but modified the wording to suit their particular needs. Items prepared in this way were generally recorded as 'teacher-made'.

The extent of the use of 'teacher-made' items is shown in Table 13. In general, a large proportion of the items are 'teacher-made', particularly the essays and assignments. The only types of items which are extensively 'borrowed' are the multiple choice and completion/short answer items, and the extended response problems in Maths I. The extent of using 'borrowed' items affects the validity and reliability of assessments and this is discussed later in the fourth section of this chapter.

Table 13: Type of item by variation in percentage of teacher-made items in the summative assessment programs

Type of Item		Minimum	Maximum	Median	Mean
(a) ENGLISH					
Multiple-choice		0	60	0	16
Completion (short answer)		0	100	50	47
True/false		0	0	0	0
Extended response (Essay/maths problem)		40	100	90	81
Assignment	In class	80	100	100	94
	At home	80	100	100	97
	Group	0	0	0	0
Practical work/oral		100	100	100	100
(b) MATHS I					
Multiple-choice		0	0	0	0
Completion (short answer)		33	50	50	44
True/false		0	0	0	0
Extended response (Essay/maths problem)		0	90	50	47
Assignment	In class	0	0	0	0
	At home	0	0	0	0
	Group	0	0	0	0
Practical work/oral		0	0	0	0
(c) FOREIGN LANGUAGES					
Multiple-choice		0	100	10	44
Completion (short answer)		0	100	100	71
True/false		0	100	0	0
Extended response (Essay/maths problem)		25	100	100	78
Assignment	In class	0	60	0	0
	At home	0	100	0	0
	Group	0	0	0	0
Practical work/oral		0	100	100	88

Use of common testing within schools

The percentage of the total assessment program that was common across the Year 12 English classes within a school varied from 29 per cent to 97 per cent (see Table 14). However, most schools used more than 70 per cent common assessment in English. Inspectors felt that in English there should be some common and some different assessment between classes. One inspector felt that more schools would be making use of different assessments between classes with the introduction of the new English syllabus. In contrast, the Maths I teachers interviewed always used common assessment wherever two or more Year 12 classes were studying the same unit of mathematics. (Inspectors recommended the use of common assessment in mathematics.) No school in the sample had more than one class studying the same foreign language at Year 12 level; therefore, this issue was not relevant for foreign languages.

Table 14: Percent of assessment that is common across classes by subject

	Minimum	Maximum	Median	Mean
English	29	97	77	71
Maths I	100	100	100	100
Foreign Languages	Not applicable			

Decision making for summative assessment

Principals and teachers were asked who made decisions, and whether they would like any changes in the decisions made with respect to:

- when tests are administered
- time allowed for tests
- weighting given to assessment elements (tests, assignments, etc.)
- mode of assessment used
- type of items used
- setting of assessment instruments
- use of common testing (within and between schools)

Most decisions were made by the subject masters, usually in consultation with the other teachers. In all cases, the principal could over-rule the decisions made by subject masters, but it appeared that this power was seldom exercised. The pattern of decision making thought desirable by inspectors corresponded well with what appears to be occurring in schools.

The areas in which the principal made decisions were in the timetabling of mid- and end of semester tests and in setting limits on the length of formal tests, and, in some cases, the maximum number of words required in assignments. In these areas, subject masters were usually consulted before decisions were made.

When tests are administered and time allowed

Table 15 shows that few principals or teachers wanted changes in the timing of the assessment instruments or in the time allowed for each of these.

Most of those who wanted to change when tests were administered either wanted to delete the mid-semester test, or make it closer to the holidays.

Of those who wanted to change the time allowed for tests written under examination conditions, three teachers wanted an increase, and two principals wanted a decrease. The change desired by the other teacher could not be classified as a decrease or as an increase. The three teachers wanting an increase were presently restricted by a maximum of 60 or 90 minutes. One of the principals desiring a decrease wanted to avoid the use of three hour tests as he considered these too long. The other principal wanted to decrease the total testing time. (As he had already imposed a 60 minute maximum on the length of tests, this could mean having fewer tests.) It is interesting to note that one of the teachers wanting an increase and one of the principals wanting a decrease in the time allowed for testing were in the same school.

Table 15: Number of principals and teachers* who wanted changes in the decisions relating to summative assessment

		ENGLISH		MATHS I			
	PRINCIPAL (12)	SAC** (12)	'Other Teacher' (9)	SAC (12)	'Other Teacher' (8)	FOREIGN LANGUAGE SAC (11)	TOTAL (64)
When administered	3	1	1	2	1	2	10
Time allowed	2	0	1	1	1	1	6

* One or more responses were allowed. Therefore the number of schools in which each category of respondent was interviewed, is indicated in brackets.

** Subject Area Co-ordinator

Since the introduction of internal assessment, there has been considerable criticism from many sectors that the assessment load for students has been excessive. Evidence relating to these criticisms has been documented in the report by Campbell et al and the Schools Under Radford Study. The data reported in these studies relate to the system as it operated in 1974.

The First Interim Report of the Select Committee on Education in Queensland recommended that the Board ... *take steps to ensure that the importance of avoiding over-testing is regularly brought to the attention of school inspectors and school principals.*

However, the evidence collected in 1978 in the present study suggests that principals are acutely aware of this problem, and that many are taking active steps to ensure that over-testing does not occur. The evidence given by students and teachers indicates that they do not consider over-testing to be a serious problem in the subjects investigated here.

Although this study did not specifically probe other subjects, informal discussions with students suggest that students studying mainly humanities subjects could possibly be experiencing an excessive assessment burden, while students in maths-science courses appear to have reasonable assessment loads.

Weighting on various assessment elements, modes of assessment and types of items

Although the subject area co-ordinators (SACs) generally made the decisions in these areas, the 'other teachers' were usually consulted before decisions were made.

While the majority of principals and teachers were satisfied with the present weightings given to the elements of assessment, modes of assessment and types of items, about one-third were dissatisfied with each area (see Table 16).

Table 16: Number of principals and teachers* who wanted changes in the decisions relating to weightings

WEIGHTING ON:	PRINCIPAL (12)	ENGLISH		MATHS I		FOREIGN LANGUAGE SAC (11)	TOTAL (64)
		SAC ** (12)	'Other Teacher' (9)	SAC (12)	'Other Teacher' (8)		
Elements of assessment	0	7	1	1	2	5	16
Mode of assessment	1	7	4	2	1	5	20
Type of item	2	5	2	2	3	4	18

---

* One or more responses were allowed. Therefore the number of schools in which each category of respondent was interviewed, is indicated in brackets.

** Subject Area Co-ordinator

Only one of the 12 principals indicated a desire to change any of these aspects of assessment. This principal currently limits the use of multiple choice items, and wishes to further limit their use.

About half of the English subject area co-ordinators wanted an increase in weighting of oral/aural assessment and a corresponding decrease in the weighting given to literature. However, several respondents doubted the validity of oral assessments. 'Other teachers' indicated that they did not want changes in the weightings presently used.

Three of the 21 English teachers wanted to broaden the types of items used, including more multiple choice items if good item banks were available. Such item banks would make multiple choice items more acceptable at moderation meetings. Two other respondents wanted more variety in the practical and oral aspects in English. They wished to include interviewing, filming and television work. One respondent indicated that items should be better matched to objectives. Two respondents favoured increasing 'at home' assignments as these developed skills which would be useful in later life.

Most of the 20 Maths I teachers were reasonably satisfied with the present weightings on elements of the course, mode of assessment and types of items used. These respondents indicated that they would like a small weighting (up to 10 per cent) for an assignment, while a further three wanted a few multiple choice items. These teachers recognised the need for the development and validation of this type of item. Two respondents wanted to use more short-answer questions, while another felt that open book examinations were a more realistic preparation for further education and for real life. This teacher felt that she would have difficulty justifying open book examinations to moderation meetings.

About half of the 11 foreign language teachers expressed a desire for change in these areas. Four of the respondents wanted an increased emphasis on the cultural aspects of the course. Others supported the view expressed by Maths I teachers desiring standardised items. Some teachers expressed concern that moderation meetings based decisions about ratings on a subset of the syllabus which was easy to examine. There was no pattern in the remaining comments as the teachers often expressed contrary opinions.

Setting of assessment instruments

Within the three subject areas in the 12 sample schools, a total of 35 assessment programs were studied. In 33 of these, assessment instruments were set by one or more of the group of Year 12 teachers in that subject. This always included the subject master where he taught a Year 12 class. In the other two cases the subject master set the examination alone even though he did not teach a Year 12 class. It appeared that, in most schools, teachers of the Year 12 classes either had an input in the setting of tests or else had the opportunity to peruse tests before their administration.

In all cases the subject area co-ordinators decided who prepared each assessment instrument. They used a wide range of criteria in determining who should prepare these instruments.

In 31 of the 35 cases, respondents indicated that they were satisfied with the allocation of who sets the tests. Four subject area co-ordinators would increase the degree of delegation of test setting if it were not that they were the one held accountable for the quality of the school's assessment program in that subject.

Use of common testing (within and between schools)

Only one of the 12 principals wanted to change the current practice relating to common tests between classes in his school (see Table 17). This principal wanted to increase the proportion of different tests. About half the other principals allowed some different testing between classes in some subjects. In schools which had all common tests between classes, the principals expressed no desire to introduce different tests. They felt that common tests were necessary to achieve comparability.

Table 17: Number of principals and teachers* who wanted changes in the decisions relating to common/different tests

		PRINCIPAL (12)	ENGLISH		MATHS I		FOREIGN LANGUAGE SAC (11)	TOTAL (64)
			SAC (12)	'Other Teacher' (9)	SAC (12)	'Other Teacher' (8)		
Common/ Different Tests	** Within schools	1	2	1½ ***	0	0	0	4½
	Between schools	6	5	3	8	5	8	35

* One or more responses were allowed. Therefore the number of schools in which each category of respondent was interviewed, is indicated in brackets.

** This question is inapplicable where there is only one class in that subject in the school.

*** In one school, two English 'other teachers' were interviewed and they held opposing views on this issue.

In two schools, both the subject area co-ordinators and 'other teachers' of English wanted to decrease the weight given to common tests between classes. Most of the schools not wanting a change already had different assessments between classes, ranging from 20 to 67 per cent.

In Maths I, all schools were using all common testing between classes. No respondent desired a change. There appeared to be a general feeling that this was the only appropriate procedure for Maths I.

In foreign languages, the issue of common testing between classes was inapplicable since each school had at most one Year 12 class in each foreign language.

Half of the 12 principals opposed any form of common testing between schools. They felt present syllabuses allowed flexibility which would be lost with either external examinations or reference tests because teachers would focus on the examinable content of the syllabuses. Of those wanting common testing between schools, two wanted ratings to be based entirely on external assessment, two wanted external assessment of a core containing about half of the syllabus with the remainder to be assessed internally in the schools, and two wanted internal assessment with reference tests to provide a guide for teachers.

None of the 21 English respondents wanted a return to full external examinations, although two subject area co-ordinators wanted external assessment of a core curriculum. A further two respondents wanted reference tests, while four wanted informal sharing of tests and items with nearby schools.

Of the 20 Maths I teachers, eight respondents wanted a full external examination with a further two wanting external assessment of a core curriculum. One respondent wanted a reference test, while a further two wanted informal sharing of tests and items. The remaining seven did not want any common testing between schools.

In foreign languages, three of the eleven respondents wanted the core curriculum tested externally, three wanted reference tests and two wanted informal sharing of test materials. The other three respondents did not want any common testing between schools.

There was a sharp contrast between the Maths I teachers, many of whom wanted a return to an external Senior-type examination, and the English and foreign language teachers, none of whom wanted a return to full external examinations.

(3) FORMATIVE ASSESSMENT

Formative assessment includes all tests, assignments, practical and other work which provides feedback on a student's rate of progress and areas of weakness.

The following aspects of formative assessment are considered in this section:

- the use of summative assessment for formative purposes;
- the extent of use of formative assessment; and
- the balance between summative and formative assessment.

Use of summative assessment for formative purposes

Return of marked papers

Principals, subject area co-ordinators, 'other teachers' and students generally agreed that test papers were given back to students after tests in all three subjects studied. In one school where the English course was divided into units, with each teacher teaching a particular unit to each class in turn, assignments were not returned to students until the end of the semester.

The only other case where papers were not returned to students involved end of semester examinations in semester four. In schools where these were held late in the year, students had left school before these had been marked.

Inspectors unanimously agreed that test papers should always be returned to the student as appears generally to be the case in schools.

Delays

All teachers and students agreed that the delay in getting Maths I and foreign language papers back was satisfactory. In general, for both these subjects papers were returned within one week (see Table 18). Allowing for the fact that delays could differ from class to class in Maths I and that the responses were generalised across many instances of assessment, there was close agreement amongst subject area co-ordinators, 'other teachers' and students as to the length of the delay in a particular school.

In English, there was reasonable agreement on the length of the delay in all except one school. (In that school, the delay perceived by teachers was up to a week where as the students perceived the delay as being greater than two weeks.) In English, the delays were longer than in the other two subjects. The delay in English in receiving papers and assignments back was considered unsatisfactory by students in four schools, while in two other schools student feeling on this issue was divided. In the other six schools, English students were satisfied with the length of the delay. In contrast, almost all English teachers perceived the delay as being satisfactory.

Table 18: Teacher perceptions of the delay in returning marked assessments by subject

LENGTH OF DELAY	ENGLISH (21)	MATHS I (20)	FOREIGN LANGUAGES (11)
1–2 days	1	5	4
3–7 days	13	13	7
8–14 days	6	2	0
> 14 days	1	0	0

Six principals felt that marked papers should be returned to students within a week, while five others felt that the maximum acceptable delay was two weeks. The other principal perceived a delay of greater than two weeks as being acceptable. In contrast with most teachers, the majority of inspectors felt that delays of between one and three weeks were acceptable.

The above results provide a sharp contrast to that reported in the Schools Under Radford Study. The findings reported in the study indicate that in 1974

> *... the use of assessment for summative purposes and the need for comparability of ratings have almost eliminated its value as feedback. This is due in part to the long delays caused by moderation requirements within the school. This is supported by the 57 per cent of teachers who reported that students sometimes experienced long delays ... and 55 per cent of grade 12 students reported that, overall, the delays were too long. Students reported that delays of several weeks were quite common before they received their marks, and in many cases, the scripts were never returned.*

(Schools Under Radford Study, p 22)

By contrast, the present study indicates that scripts are always returned to students at the time they were given their marks except after semester four as noted above. Another major difference which has emerged is that the delays in receiving scripts and marks appears to have become much shorter. However, valid comparisons cannot be made because the present study focussed specifically on only three subject areas at Year 12 level, whereas the Schools Under Radford Study had a broader focus of both levels and subjects.

The extent of satisfaction with the delays has also changed considerably. In the Schools Under Radford Study, 57 per cent of teachers and 55 per cent of Year 12 students considered that the existing delays were unsatisfactory. In the present study, most of the teachers considered that present delays are satisfactory. The students in Maths I and foreign languages are also satisfied with the present delays but about half the students are dissatisfied with the delays in English.

Student attitudes

Only a few students indicated that they were interested in marks only rather than where and why they lost marks. However, teachers felt that students were more interested in marks only than in reasons for gaining or losing marks with the view to improving their performance. The principals were equally divided in their opinions on this issue.

Corrections

Students and teachers were asked whether papers were worked through in class after testing had occurred. There was a high degree of agreement between students and teachers on this point, the opinion being that papers are worked through almost always in Maths I and most of the time in English and foreign languages. The depth of discussion of worked papers varied according to such criteria as the frequency of errors of the class as a whole on some section and nature of the test items (e.g. essays were discussed only in broad terms, while a mathematics problem could be discussed in detail). The inspectors' and principals' recommendations on the working through of papers in class appeared to correspond closely with present practices in the classroom.

In some cases, students were allowed to make corrections on their papers while in other cases, the papers were kept 'clean' for moderation meetings. This seemed to depend on the individual teacher rather than the school or the subject. In quite a few cases, the students and the teachers in the same subject within the same school disagreed in their answer to this. However, this could be because the students interviewed may not have been taught by the teachers interviewed. Almost all principals, teachers and students felt that corrections could be made just as easily on another sheet of paper if they were desired. Therefore, this was not an issue of any real concern to any of these groups. All of the inspectors felt that students should be allowed to make corrections on their papers if they so desired.

Feedback to parents

There was wide variation in responses on the issue of taking papers home. Some schools had definite policies that papers must be taken home and shown to parents, while most schools discouraged this because of the need to keep papers secure for moderation purposes. Where discouragement of taking papers home occurred, the degree of discouragement varied widely. The inspectors were unanimous in commenting that students should always be allowed to take their test papers and assignments home.

Extent of use of formative assessment

Most of the teachers interviewed use formative assessment, some as often as once per week. Table 19 shows the number of teachers who use some formative assessment.

Table. 19: Use of formative assessment by teachers*

Subject		Used	Not Used
English	(21)	18	3
Maths I	(20)	14	6
Foreign languages	(11)	10	1
TOTAL	(52)	42	10

* Numbers in brackets indicate the number of teachers interviewed.

These data suggest a marked shift from the situation in 1974.

> *One of the most disturbing features of the Schools Under Radford study was the admission by teachers of the low priority given to the use of assessment to amend teaching strategies and diagnose student weaknesses.*
>
> (Schools Under Radford Study, p 22)

By contrast, in 1978, the majority of teachers interviewed reported that they included formative assessment in their teaching program.

Balance between formative and summative assessment

Summative assessment

The majority of inspectors, principals and teachers indicated that they were satisfied with the present frequency of summative assessments (see Table 20). Ten of the 11 teachers who were not satisfied with the present test frequency wanted a small decrease in the number of tests. Similarly, all of the inspectors and principals desiring a change were wanting a decrease. Overall, about one-third of each level of respondent desired a decrease in the amount of summative assessment.

Table 20: Desires of inspector, principals and teachers* regarding the amount of summative assessment

			TEACHERS			
	INSPECTORS (6)	**PRINCIPALS (12)**	**English (21)**	**Maths I (20)**	**Foreign Languages (11)**	**SUBTOTAL (TEACHERS) (52)**
Increase	0	0	1	0	0	1
Remain the same	4	8	14	19	8	41
Decrease	2	4	6	1	3	10

* The numbers in brackets indicate the number of each type of respondent interviewed.

Overall, the students were satisfied with both the number of tests and the overall summative assessment load (see Table 21). The students who wanted more tests were probed further. These were mainly Maths I students. They appeared to want the assessment load spread more evenly through the semester, rather than an increase in the total testing time, i.e. more tests in the same time.

Table 21: Students' desires regarding the number of summative assessments by subject

	English (12)	**Maths I (12)**	**Foreign Languages (11)***
Increase	0	4	1
Remain the same	11	8	9
Decrease	1	0	1

* One of the Year 12 classes interviewed did NOT contain any students studying a foreign language.

While students are reasonably satisfied with the summative assessment load in these three subjects, it is the feeling of the researchers that this is not necessarily true of all subjects. Some informal comments by students indicated a heavy workload for those students doing mainly humanities subjects.

The number of occasions in which summative assessment was administered in English and foreign languages was considerably greater than for Maths I. This is shown in Table 22.

Table 22: Variation in the number of summative assessments by subject

Subject	Minimum	Maximum	Median	Mean
English	5	15	8	8
Maths I	1	4	2	2
Foreign languages	3	11	6	7

Although in most schools only two tests were administered in Maths I in the semester, the total testing time was comparable with that in English and foreign languages (see Table 23).

Table 23: Mean testing time and assignment load by subject

Subject	Mean testing time (in minutes)	Mean number of assignments
English	250	4
Maths I	200	0
Foreign languages	270	1

These results are similar to those reported in the Schools Under Radford Study. In that study, the majority of both teachers and students reported that the number of summative assessments was adequate, with teachers tending to want fewer tests and students tending to want more tests.

Formative Assessment

Teachers interviewed in the present study generally did not see the summative assessment program as dominating their teaching, and most indicated that they were happy with the present emphasis placed on summative and formative assessment. This contrasts with the position reported in 1974.

> *Many teachers reported that the increase in time spent on test preparation (c.f. external exams) had led to a decrease in the extent to which assessments were used as diagnostic tools.*
>
> (Schools Under Radford Study, p 16)

In the present study, half of the principals and all but one of the inspectors wanted to increase the amount of formative assessment.

While most teachers were reasonably happy about this present emphasis on assessment, many expressed concern at the preoccupation of students with marks and ratings.

Table 24: Desires of inspectors, principals and teachers* regarding the amount of formative assessment

			TEACHERS			
	INSPECTORS (6)	PRINCIPALS (12)	English (21)	Maths I (20)	Foreign Languages (11)	SUBTOTAL (TEACHERS) (52)
Increase	5	6	10	8	3	21
Remain the same	1	6	10	12	8	30
Decrease	0	0	1	0	0	1

* See comment Table 20.

Most students were also satisfied about the amount of formative assessment, except in Maths I, where a majority of students in eight of the 12 schools wanted an increase.

Table 25: Students' desires regarding the number of formative assessments by subject

	English (12)	Maths I (12)	Foreign Languages (11)*
Increase	0	8	1
Remain the same	11	4	9
Decrease	1	0	0

* See comment Table 21.

In foreign languages, both students and teachers were satisfied with the number of formative assessments while both groups wanted an increase in the number of formative assessments in Maths I in quite a few schools. The only real difference that emerges between student and teacher opinion on formative assessment is in English where the students do not want a change but where teachers desire an increase in formative assessment. The reluctance of the students to have this increase probably results from the high frequency of summative testing in English.

About 30 per cent of English and foreign language teachers wanted a small decrease in summative assessment so that an increase in formative assessment could occur. In addition, some teachers (mainly Maths I teachers) wanted to increase formative assessment without changing the amount of summative assessment.

Those teachers who wanted a change in the emphasis in either summative or formative assessment were asked for their reasons, how the change could be achieved and what the effects would be. The most common reason given related to the teachers' personal philosophies about education. Many of these teachers thought that the major purpose of testing should be to provide them and students with feedback and that providing information for outside bodies (e.g. employers and tertiary education institutions) was of lesser importance.

Most of those who wanted a change in emphasis did not know how it could be achieved. Although the extent to which they could use formative assessment would appear to be largely under the control of the individual teachers, they often claimed that they didn't have time. This was attributed to the attitudes of students to assessments which don't count, overcrowded syllabus in some semester units, and, to a small extent, the summative assessment load.

(4) MARKING AND VALIDATION PROCEDURES

Marking is the link between the summative assessment program and measures of achievement. Thus the quality marking procedures affects the accuracy of the achievement measures. This section deals firstly with the marking procedures and the satisfaction of the respondents with these, and secondly, with the use of validation procedures in both the setting and marking of tests.

Marking procedures

Allocation of marking

In 10 schools, the class teacher marked the Maths I papers of his own class while in the other two schools the test setter marked the Maths I papers across all classes. While two of the three inspectors interviewed in the mathematics area preferred that each teacher marked his own class, the other preferred that all students' work be marked by the teacher who set that paper or section.

In all but one of the schools with more than one English class at Year 12 level, each teacher does part of the marking across all classes - either each teacher marks one section of a paper, or the teachers take turns in marking the papers.

As would be expected, principals commented that the method of allocation of marking varied within their school from subject to subject.

Use of marking schemes

Almost all Maths I and foreign language teachers used detailed, prepared marking schemes for all assessments. Two inspectors wanted detailed marking schemes in Maths I while another inspector did not. Most English teachers had a detailed marking scheme for such things as comprehension while essay marks were usually assigned according to a global perception of the standard of the work. One school, however, did have quite specific criteria for essay marking. Although two of the three inspectors in the English area wanted specific criteria for marking English essays, the other preferred the global perception method. While most principals seemed aware of the procedures used in various subjects in their school, three of the 12 principals indicated that they did not know whether detailed marking schemes were used or not.

Even though detailed marking schemes were not generally used in essay marking in English, teachers felt that comparability between classes was still achieved by either 'cross-marking' where each teacher marks some assessment across all classes, or by informal perusal and discussion of each other's marking. Although English teachers were confident that this latter procedure worked effectively in achieving comparability between classes, in most cases it merely took the form of casual staffroom conversation. Therefore, the researchers doubt the effectiveness of this procedure in such cases. In addition, in most schools, the 'cross-marking' procedure did not apply to all sections of work. Hence, the class teachers often marked some sections for his own class (especially where there were different tests between classes). Consequently, students in a number of schools felt that comparability between classes was not achieved in English. This is shown in Table 26.

Table 26: Student satisfaction with comparability between classes by subject

	English (12)	Maths I (12)	Foreign Languages (11)
Satisfied	5	2	0
Student opinion divided	1	2	0
Dissatisfied	3	1	0
Not applicable (< two classes)	3	7	12

Most Maths I teachers felt that the use of detailed marking schemes ensured comparability between classes and consequently that perusal of each other's papers was unnecessary. Students did not perceive any inequities resulting from each teacher marking his own class using a detailed marking scheme. The apparent student dissatisfaction with comparability in Maths I (see Table 26) was found, on further probing, to be the result of three components - different units being taught to different classes within the school in the same semester, vertical streaming of classes doing the same unit, and the quality of teaching. The last two of these relate to the teaching process rather than to assessment. The issue of different units appears to be a problem only in schools either where the unit was being taught for the first time and teachers were unfamiliar with it, or where the students doing different units were of markedly different ability levels.

Comparability between classes

The issue of comparability between classes is not relevant in foreign languages, since no school had more than one class in the same foreign language at Year 12 level.

Most principals and teachers did not consider the achievement of comparability between classes to present any problems. This opinion was generally shared by the inspectors. However, two of the six inspectors feel that comparability between schools is not being achieved.

This confidence in the achievement of comparability between classes was not shared by Elley & Livingstone (1972). They were confident of teachers' abilities to rank students in their class, but expressed concern about their ability to compare performance between classes or between schools.

> *Most teachers have had experience in testing and arranging pupils in rank order according to their attainment within a class. A more difficult task is that of relating the performance of one class to that of others in such a way that similar standards (i.e. levels of attainment) receive similar assessments. This problem of equating standards across schools and subjects has proved to be one of the main stumbling blocks to internal assessment in other countries ...*
>
> (Elley & Livingstone, p 45)

Consistency of marking from year to year

Almost all teachers felt that their marking standards were consistent from year to year. Teachers felt that they achieved this through a combination of subjectivity and piecemeal re-use of some test items. Amongst the subjective criteria were staff stability, teaching experience and teacher perception of the quality of work necessary to achieve a particular rating. None of the procedures used were well organised nor were they based on educational measurement techniques.

Where decisions were required concerning marking schemes or the use of 'cross-marking', such decisions were made by the subject master. Other teachers were generally satisfied with the decisions made.

Validity and reliability of assessment instruments

Most principals and teachers (56 out of the 64) indicated that they were satisfied with the quality of the assessment instruments in their school or subject. Respondents were then questioned further on how validity and reliability of tests were obtained. Most indicated that their view was based purely on their own perceptions rather than on any specific validation procedure. However, some teachers commented that procedures used in some of the 35 assessment programs investigated (e.g. co-operative test setting, 'cross-marking', and the use of detailed marking schemes) contribute towards validity and reliability even though that may not have been the reason for their use.

Brown (1976, p 51) comments that *a test must possess two characteristics if it is to be considered a good measuring instrument: the test must be reliable and it must be valid.* These terms are defined as follows in the Australian Item Bank, Handbook for Mathematics Item Bank (Australian Council for Educational Research, 1978, p 15).

> *Although there are different types of 'validity', probably the most relevant for classroom teachers is content validity - that is, the extent to which a test item measures performance in the subject matter of the course that has been taught. To determine whether items are valid for a particular group, the educational objective being tested by each item should be examined. Teachers must make their own judgments on the relevance of items for their students.*
>
> *Like validity, the term 'reliability' has several interpretations. One is that reliability is a measure of consistency over a period of time; that is, if the same test is administered to the same students some time later, would the rank order of students be unchanged? Another interpretation is that of internal consistency - that is, are all the items in the test measuring the same broad objectives?*
>
> *Both validity and reliability of tests may be affected by items which are either too difficult or too easy for the students being tested, by items which do not discriminate effectively, and by items which have unsuitable distractors.*
>
> (Australian Item Bank, Handbook for Mathematics Item Bank - Australian Council for Educational Research, 1978, p 15)

In contrast to the apparent satisfaction of respondents in schools with the quality of tests, only one of the six inspectors interviewed was satisfied with their quality. Another three felt that the quality of tests varied markedly, while the remaining two inspectors were generally dissatisfied with test quality.

When asked whether they felt that teachers possessed the necessary skills to ensure that tests were valid and reliable, two inspectors said that teachers did not, and the other four inspectors felt that while some teachers possessed these skills, others did not. Some of the inspectors' suggestions for improving the validity and reliability of tests are listed below (with numbers in brackets indicating the number of inspectors suggesting each procedure).

- Relevant teachers should discuss the quality of their tests. (2)
- Principal and/or subject master should review tests to check their quality. (2)
- Improve pre-service training in assessment procedures. (2)
- Run in-service courses on assessment procedures within the school. (1)
- Seek help from some testing 'expert' outside the school. (1)
- Write down the objectives of the teaching program and ensure that the assessment tests these. (1)
- Vary the type of item used. (1)
- Spend more time on the setting of each test. (1)
- Make more item banks, etc. available to teachers. (1)
- Set tests at the beginning of the course. (1)
- Change assessment program if necessary during the semester. (1)

The views expressed here are supported by Elley & Livingstone. They noted that:

> ... *internal assessment without external examination will stretch the competence of teachers in testing and observing, in marking and scaling, in judging and interpreting. While many teachers are competent to do the work required, it is clear that much more will need to be done to assist the professional growth of all teachers in these respects. In-service training may provide some of this instruction, but sound assessment techniques cannot be mastered in a week's course. Pre-service programmes in the teachers' colleges may need to be strengthened; advisory services introduced; model tests prepared; item banks made available; clerical and computer facilities investigated, and other means examined for helping teachers with their assessments ...*
>
> (Elley & Livingstone, p 45)

(5) ACHIEVEMENT AND OTHER EXIT MEASURES

Data was collected relating to the measuring and reporting of student characteristics for Years 11 and 12. The measures provided about the students are:

- moderated school assessments
- special subject assessments
- school reports.

The moderated school assessment is a rating of a seven-point scale in each subject for each semester. After approval by the Board of Secondary School Studies, these appear on the Senior certificate issued at the end of Year 12.

The special subject assessment is a score on a 99-point scale assigned by the school in each subject. This score reflects the schools' assessment of each student's overall performance in each subject over all semesters in which the subject is studied.

The school report contains information in each subject and is issued by the school at least once per semester. The nature of the report varies considerably among schools, depending on individual school philosophies.

Moderated school assessments

The abolition of external examinations with the introduction of internal assessment gave schools some freedom in the methods used in assessing student performance. Student ratings are now based on a variety of types of school assessments which are administered on various occasions as determined by the school, instead of being determined by three to six hours of formal written examinations administered on one day after the completion of the course. Information on the Board's certificates is contained in Appendix 8.

The structure of the moderation procedures is presented diagrammatically in Unit 3 of the Information Kit published by the Board of Secondary School Studies. This diagram is reproduced in Figure 2. The figure indicates the responsibilities of those involved - the school, the moderation meetings, the district moderator and the chief moderator.

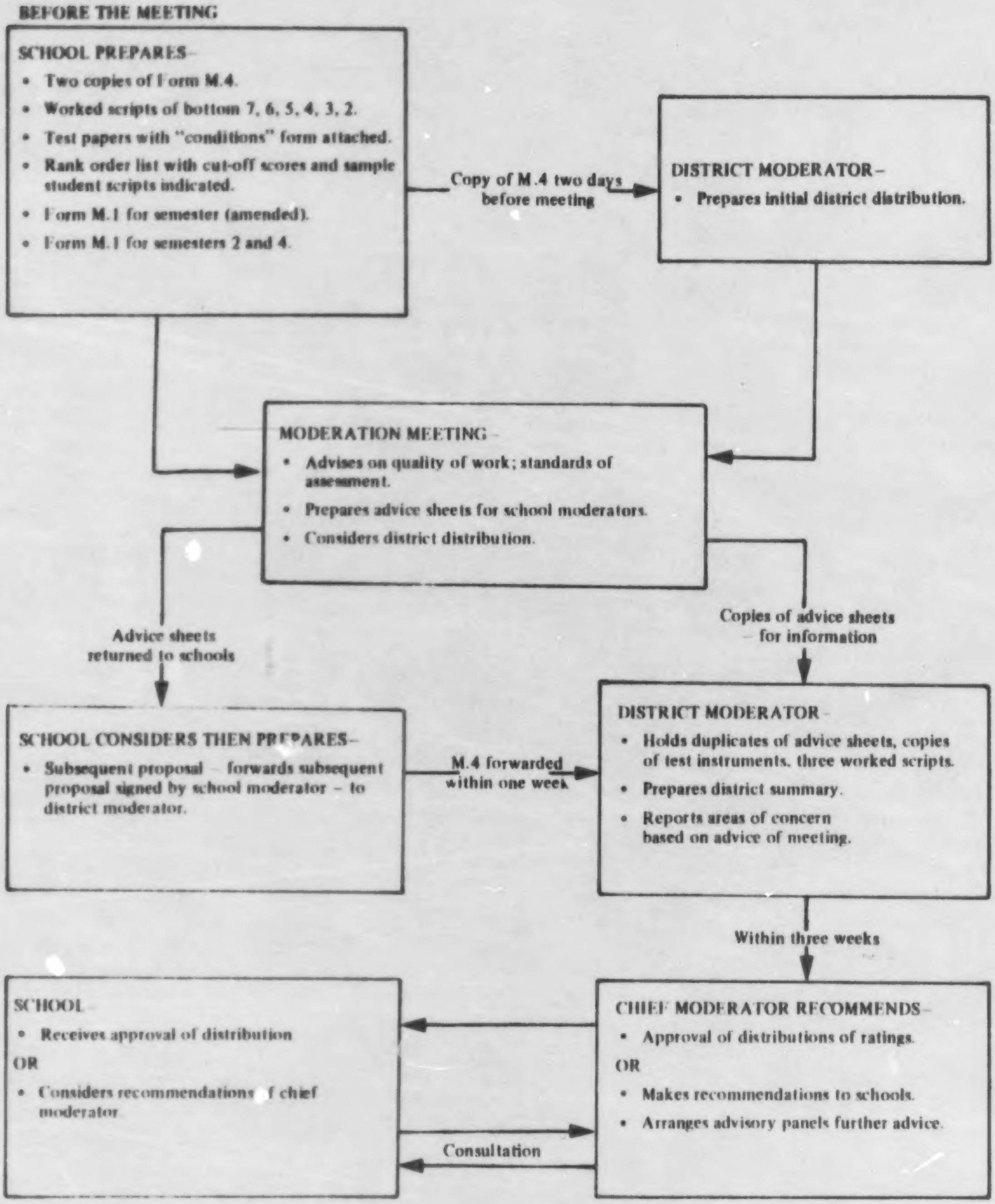
BEFORE THE MEETING
SCHOOL PREPARES–
• Two copies of Form M.4.
• Worked scripts of bottom 7, 6, 5, 4, 3, 2.
• Test papers with "conditions" form attached.
• Rank order list with cut-off scores and sample student scripts indicated.
• Form M.1 for semester (amended).
• Form M.1 for semesters 2 and 4.
Copy of M.4 two days before meeting
DISTRICT MODERATOR–
• Prepares initial district distribution.
MODERATION MEETING–
• Advises on quality of work; standards of assessment.
• Prepares advice sheets for school moderators.
• Considers district distribution.
Advice sheets returned to schools
Copies of advice sheets – for information
SCHOOL CONSIDERS THEN PREPARES–
• Subsequent proposal – forwards subsequent proposal signed by school moderator – to district moderator.
M.4 forwarded within one week
DISTRICT MODERATOR–
• Holds duplicates of advice sheets, copies of test instruments, three worked scripts.
• Prepares district summary.
• Reports areas of concern based on advice of meeting.
Within three weeks
SCHOOL–
• Receives approval of distribution
OR
• Considers recommendations of chief moderator.
CHIEF MODERATOR RECOMMENDS–
• Approval of distributions of ratings.
OR
• Makes recommendations to schools.
• Arranges advisory panels further advice.
Consultation

Figure 2: Sequence of events in the Board's moderation procedures

Derivation of final scores

Current assessment procedures necessitate the combining of test scores by schools to obtain a single score before moderation. Under the moderation system, schools are required to specify the weightings they intend to give to each element of their assessment programs. It can be shown empirically that to achieve the desired weightings, it is necessary to allow for differences in standard deviations between tests when combining test scores. This point was made in the School Assessment Procedures Booklet, Number 4, *Moderation within schools* (Department of Education, Queensland, 1972, p 38). This series of booklets was prepared to guide teachers developing strategies to assist in implementing the moderation system. In the subjects investigated in the 12 sample schools, all respondents indicated that the final score was obtained by merely adding raw scores. Although exact calculation of standardised scores can be laborious, there are several simple methods of approximation available. However, teachers do not appear to be using any of these techniques and appear unaware of the effect of standard deviation of test scores on weightings. In contrast, all of the inspectors were aware of the effect of differences in standard deviations and all but one inspector recommended that standardised scores be added rather than raw scores.

Determination of student ratings

In all cases, the responsibility for proposing cut-off marks on the Board of Secondary School Studies' M4 form rested with the subject area co-ordinator. (The cut-off mark for a particular rating is the minimum student score which the school believes is necessary to obtain that particular rating.) These proposed cut-offs had to be approved by the school moderator (principal) before being submitted to moderation meetings. In three of the nine schools with more than one English teacher involved at Year 12, the subject area co-ordinator consulted the other teacher involved during the determination of cut-offs. In six of the eight schools with more than one Year 12 Maths I class, the other teacher was also involved in determining cut-offs. This issue was not applicable in foreign languages.

Respondents were asked what criteria they used in deciding on appropriate cut-offs. Many respondents applied several criteria. The results are shown in Table 27.

Table 27: Criteria used in determining cut-off marks on the M4 form

CRITERIA	PRINCIPAL (12)	ENGLISH		MATHS I		FOREIGN LANGUAGE SAC (11)	TOTAL (64)
		SAC * (12)	'Other Teacher' (9)	SAC (12)	'Other Teacher' (8)		
Fitting the normal curve	7	5	2	2	3	0	19
Cut-offs at pre-determined percentage scores	1	3	1	7	1	5	18
Teacher perception of standards	10	10	6	8	6	11	51
Distribution of these student's ratings in previous semesters	2	5	1	1	2	2	13
Distributions of ratings of last year's Year 12 students	0	2	1	0	2	0	5
Cut-offs at gaps in the distribution	2	4	1	4	4	3	18
Other	3	1	1	2	1	3	11

* Subject Area Co-ordinator

For each of the six groups of respondents in schools, the major criterion used was teacher perception of standards. Substantial numbers of respondents used:

- filling the normal curve;
- cut-offs at pre-determined percentages;
- the distributions of these students' ratings in previous semesters in the subject; and
- cutting off at gaps in the distribution of marks.

Only five of the 64 school staff interviewed used the distribution of the ratings of Year 12 students in the previous year.

All inspectors felt that teacher perception of standards should be the main criterion used to determine cut-off marks.

The moderation system

Respondents in schools were asked for their perception of the nature of comments made at moderation meetings. Table 28 shows that the positive comments outweigh the negative comments by about two to one. The only group with more negative than positive comments were the principals. Many principals and teachers indicated that non-informative comments such as 'okay' or 'satisfactory' were often made at moderation meetings. Most respondents thought that the comments made at moderation meetings were generally justified.

Table 28: Perception of the nature of comments received at moderation meetings by type of respondent*

NATURE OF COMMENT	PRINCIPAL (12)	ENGLISH		MATHS I		FOREIGN LANGUAGE SAC (11)	TOTALS	
		SAC** (12)	'Other Teacher' (9)	SAC (12)	'Other Teacher' (8)			
Useful	1	3	2	4	4	5	19	46 Positive Comments
Constructive	3	8	2	6	4	4	27	
Irrelevant	6	3	2	4	1	3	19	24 Negative Comments
Critical	0	2	1	1	0	1	5	
Other	3	1	4	3	2	1	14	

* Respondents were asked whether they frequently received each of these types of comment.

** Subject Area Co-ordinator

Respondents were asked whether the moderation meeting recommended a change in the distribution of ratings in their subject in semester three. (At the time that data was collected, moderation in semester four had not yet occurred.) Eighteen of the 64 indicated that downward movements of ratings had been recommended, while upward movements were recommended to three others. However, only about half of the changes recommended were subsequently made. Some of those who accepted downward movements still believed that the quality of the students' work justified the original proposed ratings. In some cases where teachers and administrators had different opinions on the appropriate distribution of ratings, the teacher's proposal was submitted on the understanding that the recommendation of the meeting would be the deciding factor.

Inspectors, principals and teachers were asked whether they were satisfied with present moderation procedures. Those who were dissatisfied were asked for their reasons, possible alternatives and the effects of their alternatives. They were also asked if they had suggested changes, and if so, to whom.

Table 29 shows that about two-thirds of the respondents were dissatisfied with the moderation system.

Table 29: Number of respondents satisfied with present moderation procedures

DEGREE OF SATISFACTION	INSPECTORS (6)	PRINCIPAL (12)	ENGLISH		MATHS I		FOREIGN LANGUAGE (11)	TOTAL (70)
			SAC * (12)	'Other Teacher' (9)	SAC (12)	'Other Teacher' (8)		
Satisfied	2	3	0	5	3	2	2	17
Mixed comments	0	1	1	1	3	1	1	8
Dissatisfaction	4	8	11	2	6	5	8	44
No comment	0	0	0	1	0	0	0	1

* Subject Area Co-ordinator

The changes suggested by inspectors were very diverse. One inspector suggested accreditation of schools while another was strongly opposed to that idea. Other suggestions for change were regionalisation by syllabuses; that the district moderator should visit each school; and the perusal of other schools' papers before the moderation meeting so there is more time at the meeting for discussing them.

Two of the six inspectors had suggested changes to the Board while another had made suggestions for change at a meeting of principals.

Three of the 12 principals said that they would prefer a district review panel. This panel would contain fewer members than the present moderation meetings. They felt that only teachers with sufficient experience to make competent judgements about standards of achievement in their subject should be on these panels. They also felt that this would cause less disruption to individual schools.

A further two principals indicated that they wanted to change to a system of competency-based assessment as proposed by the *Review of School-Based Assessment in Queensland Secondary Schools*. One felt that it is inappropriate for teachers to comment on the work of other teachers.

In addition, each of the following changes was suggested by a principal:

- providing a single exit measure rather than separate ratings at the completion of each semester because an exit achievement is simpler and more meaningful to employers and tertiary institutions;
- strengthening the present system by increasing the time devoted to examining scripts at moderation meetings because it is not possible to evaluate the material adequately in the time presently available;
- giving schools accreditation to provide their own ratings without the requirement of moderation because teachers are competent to make judgements about their own students; and
- return to an external examination because they are fairer and reduce the constant bombardment of assessment on students.

About half of the principals had suggested proposed changes at meetings of principals or deputy principals. Most principals were aware of some of the effects of their suggested alternatives.

While most English subject area co-ordinators were dissatisfied with the moderation system, the majority of English 'other teachers' were reasonably satisfied. Seven of the 21 respondents in English wanted an alternative which decreased the emphasis placed on moderation. The alternatives they suggested were:

- give accreditation to schools because moderation would become increasingly difficult with the new syllabus;
- eliminate moderation because it does not achieve comparability but still have meetings where teachers from different schools can swap ideas;
- replace moderation meetings with a panel of experienced teachers because this will increase comparability; and
- have moderation meetings only for small schools and schools with inexperienced teachers because it is a waste of time for experienced teachers in larger schools.

Two of the 21 respondents in English wanted an increase in the commitment to moderation. One wanted more time spent examining scripts, while the other wanted larger schools to send an additional teacher to enable that teacher to gain experience in moderation. It was felt that at present scripts were only being superficially examined because of insufficient time and manpower at moderation meetings.

Four of the 21 English respondents had suggested changes outside the school (to district moderators or principals' meetings) while two others had suggested changes to others within the school. Very few of those suggesting change explained any ramifications of their proposed change.

There was considerable dissatisfaction with present moderation procedures among both subject area co-ordinators and 'other teachers' in Maths I. As was reported in the second chapter of this report, ten of these 20 respondents wanted a return to external examinations as they believed this would be fairer than moderation and would achieve comparability. Some of these wanted internal assessment as well so that both students seeking tertiary entrance and those seeking employment would be catered for. Three respondents wanted to change to competency-based assessment, while a further three wanted to tighten up the present system, either by allowing more time for moderation, or by ensuring that the recommendations of the moderation meetings were followed. Respondents felt that where schools ignored recommendations, inequities would exist and comparability would be lost. Two respondents considered the February meeting a waste of time, while another perceived it to be the most valuable. The teachers who recommended a move to competency-based assessment commented that there was a need to implement it slowly and that its implementation would require extra work on the part of teachers.

Of the Maths I teachers suggesting changes, four respondents had discussed changes with the administration within their school, four others had suggested changes at moderation meetings, and one had made a submission to the Select Committee on Education in Queensland. Most of those who proposed a return to external examinations had a reasonable perception of the positive effects of such a move, although only one mentioned any of the negative effects.

Six of the 11 foreign language teachers wanted either a reference test or external assessment of some parts of the course, such as grammar. They felt that comparability was not being achieved in that some meetings allowed lower quality work to slip through and some participants based decisions about ratings only on the sections of the course which they considered important. Other suggestions included reference testing, informal swapping of papers, the use of item banks and more meetings of teachers for discussions instead of moderation. Several teachers felt that foreign language students were a small group of above average intelligence and therefore it was inappropriate to apply a normal curve to the distribution even within a moderation district.

Three foreign language teachers had suggested changes at moderation meetings. Those who suggested changes seemed to be aware of both the positive and negative effects of their proposals.

The implication of marking reliability on comparability

The major aim of moderation is to achieve comparability of standards among schools. Each school submits proposed distributions of ratings based on the perception of standards of the staff of that school. Agreement at moderation meetings requires that the reviewers have similar perceptions of standards to staff in the school submitting the papers. However, research does not support the notion that teachers are able to judge standards. Lewis (1974) lists several reasons why essay questions cannot be marked reliably. These include:

> ... *They* (essay questions) *are difficult to mark reliably, and with only a single marking the subjective element is very considerable. The correlation between the marks of two examiners for the same set of answers, or even between the marks of the same examiner marking the same set of answers on different occasions, is seldom sufficiently high to justify confidence.*
>
> ... *A choice of questions enables the candidates to answer in effect different papers, so the same total mark may not represent comparable performances ...*
>
> (Lewis, p 107)

Theobald (1974) makes a similar point, and cites the work of Hartog & Rhodes (1938) showing the unreliability of a marker when he marks essays on several occasions.

In each of the examples cited above, comparisons were made within a particular student group. Lewis claimed that even with the same marker, the use of different questions produced results which were not comparable. Both Lewis and Theobald asserted that even when the same questions were used, the use of different markers led to unreliable results. Further, Lewis and Hartog & Rhodes claimed that even when the same marker re-marked the same question on different occasions, that marker was unable to produce consistent results.

In each of the studies described above, it was shown that marking of essays was unreliable when only one of the parameters varied. In the moderation process, participants attempt to compare standards of work in a situation in which all parameters vary.

In view of this evidence, it is not surprising that less than one quarter of respondents (as shown in Table 29) were satisfied with present moderation procedures. It is possible that the consensus reached at moderation meetings stems from the need to make a decision rather than from a belief that comparability has been achieved.

Special Subject Assessments (SSAs)

Derivation of SSAs

In eight of the sample schools, the principals indicated that each subject area co-ordinator was responsible for preparing the Special Subject Assessment in his subject area. In the other four schools, the principals' response was that the SSA was prepared by a mathematics 'expert' working in conjunction with the administration. In two cases, this 'expert' was the mathematics subject master while in the other two cases, it was an 'other teacher' of mathematics. Table 30 shows the teachers' response to the question of who prepared the SSA in their subject. The numbers in bold type represent the responses of the principals in those schools which had this category of respondents. The difference between the numbers in the cells and the numbers in bold type indicates the number of teachers who did not know their school's policy. Of the 52 teachers (all of whom taught at Year 12 level), 11 of them did not know whose responsibility it was to prepare the Special Subject Assessment.

Table 30: Teachers' perceptions of who prepared the Special Subject Assessment (SSA) in their subject

WHO PREPARES THE SPECIAL SUBJECT ASSESSMENT?	**ENGLISH**		**MATHS I**		**FOREIGN LANGUAGE SAC (11)**
	SAC * (12)	**'Other Teacher' (9)**	**SAC (12)**	**'Other Teacher' (8)**	
SAC for *his* subject	$8^{\mathbf{8}}$	$4^{\mathbf{7}}$	$6^{\mathbf{8}}$	$4^{\mathbf{6}}$	$4^{\mathbf{7}}$
SAC and 'other teacher' jointly for *their* subject	0	0	2	2	1
A maths 'expert' in conjunction with the administration of *all* subjects	$3^{\mathbf{4}}$	$0^{\mathbf{2}}$	$4^{\mathbf{4}}$	$2^{\mathbf{2}}$	$1^{\mathbf{4}}$
Don't know	1	5	0	0	5

(The numbers in bold type indicate the responses of the principals in those schools which had that category of respondents.)

* Subject Area Co-ordinator

When asked how the Special Subject Assessment was prepared, 22 of the 52 teachers admitted that they did not know (see Table 31). Of the remaining 30, many were unable to give a clear explanation of the steps used in preparing the SSA in their subject. Many teachers took great interest in the ratings which their students received, but appeared disinterested in the assigning of SSAs and its relationship to the Tertiary Entrance (TE) Score. In view of the importance of the TE Score to the students, it was surprising to find among teachers such a lack of knowledge of who prepared the SSA, how it was prepared and its effect in determining the TE Score.

Table 31: Teachers' awareness of how the Special Subject Assessment (SSA) is prepared in their subject

DO YOU KNOW *HOW* THE SPECIAL SUBJECT ASSESSMENT IS PREPARED IN YOUR SUBJECT?	**ENGLISH**		**MATHS I**		**FOREIGN LANGUAGE SAC (11)**	**TOTAL (52)**
	SAC* (12)	**'Other Teacher' (9)**	**SAC (12)**	**'Other Teacher' (8)**		
Yes	9	1	11	5	4	30
No	3	8	1	3	7	22

* Subject Area Co-ordinator

Detailed questioning was carried out in order to determine just what steps were used in the preparing of the SSA in each school. The questions asked included the following:

- Which semesters were included and with what weightings?
- Was the starting point for the estimate raw scores, ratings or subjective judgement?
- Were SSAs adjusted in special cases and, if so, for what reason?
- Were scores standardised before input, after input, or not at all?
- Was the ASAT mean (either this year's or last year's) used to estimate TE Scores?
- Were scores re-scaled to a particular mean and standard deviation and, if so, what were these?
- Was the shape of the distribution of scores taken into account?
- Were SSAs shown to the Year 12 teachers of that subject who were not involved in preparing them?

The most noticeable feature that emerged from this questioning was the great diversity in the procedures used to calculate this score. Some schools included all four semesters work in their calculations. Others considered semester one as a 'settling in' period and therefore ignored it, while others deleted semester four because of the difficulty of finalising marks in time for them to be included in the SSA calculation. The weightings placed on the semesters used differed not only from school to school, but also from subject to subject within a school in some cases. Many schools placed a larger weighting on the later semester. Four of the six inspectors felt that this should be done.

One school used the Board approved ratings as its starting point for preparing SSAs, while the other 11 schools used the raw scores obtained in one or more semesters. All but one inspector suggested that one or more semester score be used. However, the inspectors perceived the need to standardise these before adding them if more than one of the semester scores were to be used. The remaining inspector wanted SSAs to be assigned largely subjectively but with reference to the semester scores. In some schools, after the adding of semester scores, subjective judgements were used to adjust some of these totals to allow for criteria not directly measured by the semester scores. Such adjustment was recommended by four inspectors. By contrast, some other schools stuck rigidly to the mark obtained from the combination of raw scores. All schools were prepared to adjust the SSA of an individual student in cases of personal hardship beyond his control (e.g. injury, bereavement, etc.). All inspectors agreed with this practice.

Table 32 shows that eight of the schools standardise scores at some stage during the preparation of the SSA. However, as mentioned previously, standardisation of scores is more desirable before combination to ensure that the actual weightings correspond with those desired. Standardisation of scores after combination has no effect on TE Scores because standardisation is the first step in the conversion of the SSA to a TE Score. Thus, only six of the eight schools using standardisation are using it effectively.

Table 32: When standardisation occurred in combining semester scores to obtain the Special Subject Assessment (SSA) for each subject

WHEN STANDARDISATION OCCURRED	NUMBER OF SCHOOLS
BEFORE combination only	4
Both BEFORE and AFTER combination	2
AFTER combination only	2
NOT at all	4

Seven schools used ASAT means to calculate the adjusted school assessment in a particular subject. Some of these went further to actually estimate TE Scores. Two schools adjusted to recognised means for standardised scores (62 or 50), while the remaining three schools did not adjust their scores because they felt it unnecessary since adjusting is the first step carried out by the Board when converting SSAs to TE Scores. Where schools adjusted the scores, the standard deviation was set to 12.

A few schools examined the gaps and groupings of scores (i.e. the shape of the distribution) to ensure that the differences between scores reflected actual difference in achievement. The majority of principals, however, felt that the distribution of SSAs would not differ significantly from that of the raw semester scores from which the SSAs were derived.

Although almost all of the principals felt that SSAs were shown to the Year 12 teachers after calculation, only about half of those teachers felt that this was done. All inspectors felt that the SSAs should be shown to the Year 12 teachers after calculation.

As mentioned previously, the over-riding pattern that emerges is one of diversity. However, this diversity does not conflict with Board policy which is not rigid with respect to procedures to be used in the preparation of SSAs.

Manipulation of Special Subject Assessments (SSAs)

Respondents were asked whether they thought it possible to manipulate SSAs to advantage some students. Table 33 shows that the majority of respondents felt that manipulation was possible.

Table 33: Respondents' perceptions of whether Special Subject Assessments (SSAs) can be manipulated

CAN SPECIAL SUBJECT ASSESSMENTS BE MANIPULATED?	INSPECTORS (6)	PRINCIPAL (12)	ENGLISH		MATHS I		FOREIGN LANGUAGE SAC (11)	TOTAL (70)
			SAC * (12)	'Other Teacher' (9)	SAC (12)	'Other Teacher' (8)		
Yes	5	11	9	3	10	5	4	47
No	0	0	2	0	1	1	0	4
Don't know	1	1	1	6	1	2	7	19

* Subject Area Co-ordinator

The researchers felt that few of the respondents had sufficient understanding of the processes involved to effectively manipulate the scores to produce the desired effects on the resultant TE Scores. This feeling was based on the admission by many respondents that they didn't know how to calculate SSAs together with the researchers' opinions that many of the proposed methods of manipulation would be ineffective. However, a few respondents indicated that they had a thorough understanding of the processes involved, and the researchers perceived that if they tried, their manipulation would achieve the desired result.

When asked whether they thought that scores were being manipulated, about half the respondents said that they did not know, while a quarter felt that scores were being manipulated, and the remaining quarter thought they were not being manipulated.

Fairness of Tertiary Entrance (TE) Scores

Respondents were asked what groups, if any, were unfairly affected by the assigning of SSAs and the subsequent calculation of TE Scores. Table 34 shows the results.

Table 34: Respondents' perceptions of whether any groups of students are unfairly affected by the assigning of Special Subject Assessments

ARE CERTAIN GROUPS UNFAIRLY AFFECTED?	INSPECTORS (6)	PRINCIPAL (12)	ENGLISH		MATHS I		FOREIGN LANGUAGE SAC (11)	TOTAL (70)
			SAC* (12)	'Other Teacher' (9)	SAC (12)	'Other Teacher' (8)		
Yes	4	7	5	1	7	5	4	33
No	1	2	2	1	3	1	1	11
Don't know	1	3	5	7	2	2	6	26

* Subject Area Co-ordinator

Although 33 of the 70 respondents said that some groups were unfairly affected, few of them were able to justify their answers in terms of the procedures used in calculating the TE Scores. Three areas of concern were isolated from the responses:

- the nature of the subject;
- the ability level of students studying the subject; and
- the size of the group studying the subject.

At this point, respondents often commented on subjects other than those investigated in this study.

Some respondents felt that the skills developed in their subject were not tested by ASAT. Thus, they felt that ASAT was not an appropriate scaling device. This claim was made for foreign languages, art and music.

Other respondents said that in some subjects, students were of higher than average ability, e.g. Maths II, physics and foreign languages. They felt that even though these subjects had a higher ASAT mean than other subjects, the rescaling did not do justice to all students in these groups.

Concerns about group size involved schools with small Year 12 enrolments and subjects taken by only a few students in the school, e.g. foreign languages, music and theatre. It was felt that small groups were disadvantaged.

In summary, almost half of the Year 12 teachers interviewed were unaware of how the SSA was prepared and some were even unaware of who prepared it in their subject. The majority of respondents felt that the SSA could be manipulated intentionally, while almost half felt the method of calculation of the TE Score unfairly affected some groups of students. In view of this apparent ignorance of and lack of confidence in the TE Score, it appears that further research to examine some of the pertinent issues should be carried out. Indeed, recommendation P27 of the *Review of School-Based Assessment in Queensland Secondary Schools* suggests that *the Board should continue its research into the efficacy of the Tertiary Entrance Score as a method of ranking students*. Further research in this area was also recommended in the First Interim Report of the Select Committee on Education in Queensland (page 15) and by Keeves, McBryde & Bennett (page 12). A proposal for such further research is outlined in Appendix 7

School Reports

The contents of the end of semester report varied markedly among schools. All schools provided at least one achievement measure, while one school provided four such measures. All except one school made provision for at least one affective measure. Table 35 indicates the number of schools using the various descriptors of achievement and affectivity.

Table 35: Number of schools using various descriptors on their end of semester school reports

DESCRIPTOR		NUMBER OF SCHOOLS (12)
ACHIEVEMENT	Individual test scores	1
	Per cent for semester	7
	Rating (seven point scale)	8
	Coded comment	2
	Written comment	3
AFFECTIVITY	Coded comment	9
	Written comment	2

Several schools also reported to parents after mid-semester tests. This report was either written or oral, and was usually less formal than that given at the end of the semester. The measures were usually test scores and open ended comments.

Class teachers were responsible for assigning all descriptors other than semester ratings. The ratings reported were either those recommended by the school on their M4 form, or the Board approved ratings.

About half of the respondents in schools were not happy with the information on the school report. Of these, many had suggested to the administration or other staff members at their school that changes be made. There did not appear to be a pattern among the suggested changes. Those with complex reporting systems often wanted them simplified while those with simple measures wanted to provide more detail. There appears to be a wide diversity in what is considered to be feasible for teachers to produce and appropriate to the needs of parents.

Inspectors were asked what information should appear on the school report. All suggested at least one achievement measure although there were differing opinions (as amongst teachers) as to what the measure should be. Inspectors also suggested that an affective measure appear on the report. While most schools appeared to use coded comments, most inspectors felt the use of written comments was much more desirable.

(6) ATTAINMENT OF THE PURPOSES OF ASSESSMENT

Many attempts have been made at defining a set of purposes of assessment for secondary schools. However, there is no universal acceptance of any one such set of purposes. *The Workshop Report* of the Conference, *Moderation - at the crossroads?* (McMorrow (ed.), p 35) listed five purposes of assessment which were claimed to be consistent with the recommendations of that report. The report of the *Review of School-Based Assessment in Queensland Secondary Schools* (page 11) proposed that the Board should support these same five purposes of assessment as appropriate for secondary education in Queensland.

There has been considerable comment in the mass media on the negative aspects of the system of internal assessment. The nature of these comments has been examined by Campbell et al (page 61) for the period 1971 to 1974. A tabulation of their findings revealed more negative than positive comments in that period. Negative comments continue to appear in the press on occasions.

Principals and teachers were given a copy of the five purposes of assessment and then asked to consider their testing program, moderated school assessments (Board ratings), Tertiary Entrance (TE) Scores and school reports and indicate whether these facilitated, militated against or had no effect on the achievement of these purposes of assessment (see Table 36).

For some of the objectives, a few respondents gave different answers to different aspects of the question (e.g. they distinguished between assessing learning and correcting deficiencies in objective 1, and between future education and employment in objectives 4 and 5). In these cases, both responses were recorded. Conversely, some respondents did not feel competent to answer for some aspects. Thus, the total number of responses in each of the 20 cells varies.

It can be seen from Table 36 that for 15 of the 20 questions, the majority of respondents saw the testing program and the exit achievement measures as facilitating the achievement of the purposes of assessment.

This is particularly evident for assessment objectives 3, 4 and 5. Very few teachers saw the testing program and exit achievement measures as militating against the assessment objectives. However, the results should be taken with some caution. Although the researchers emphasised to respondents the need to consider the testing program as distinct from the three exit measures in this question, it would appear that many failed to do so. For example, it is difficult to see how the testing program *per se* facilitates the achievement of objective 5 in normal situations.

Thus, in spite of the problems associated with the interpretation of the questions by respondents, the majority of school staff perceived their present practices as facilitating the achievement of the purposes of assessment. They did not support the view often espoused through the mass media that, under the present system, the testing program and exit achievement measures were militating against the achievement of these purposes.

Table 36: Perceptions of principals and teachers of the present achievement of the purposes of assessment

ASSESSMENT OBJECTIVES	TESTING PROGRAM			MODERATED SCHOOL ASSESSMENT			T E SCORE			SCHOOL REPORTS		
	Facilitate	No Effect	Militate Against	Facilitate	No Effect	Militate Against	Facilitate	No Effect	Militate Against	Facilitate	No Effect	Militate Against
1. To enable the student to assess his/her learning and to provide information which can be used to correct deficiencies in his/her learning.	**54**	10	2	25	**41**	4	3	**54**	3	**51**	12	0
2. To enable the teachers and students to modify teaching/learning programs to suit the needs of the students.	**51**	10	2	17	**42**	5	4	**52**	3	26	**41**	0
3. Ultimately to provide the student and parents with information concerning the student's achievement and potential.	**56**	10	1	**54**	11	2	**50**	11	0	**62**	2	0
4. To assist students and parents in determining future education and employment pathways.	**49**	15	1	**58**	6	1	**55**	7	0	**57**	7	1
5. To provide, for other educational institutions and employers, an indication of the suitability and readiness of the student to undertake further education and/or employment.	**54**	12	0	**54**	12	1	**55**	14	1	**50**	24	0

The number in bold type represents the largest number of respondents for that question.

(7) GENERAL CONCERNS AND SUGGESTIONS

Respondents were given the opportunity to elaborate on issues discussed earlier in the interview (if they desired) or to discuss issues not dealt with previously. The concerns expressed at this point by five or more of the 64 principals and teachers are listed below.

- There was too much formal testing and this was causing undue pressure on students.* (15)
- Comparability among schools was not being achieved. (10)
- Some teachers misrepresent the levels of achievement of their students to obtain higher ratings. (10)
- External examinations would be fairer. (10)
- The quality of tests varied considerably among schools with the assessment program at some schools being inadequate.* (6)
- Competency-based assessment should be introduced.* (5)
- Oral work is difficult to assess. (5)

The concerns expressed at this point by two or more of the six inspectors are listed below.

- There is too much formal testing.* (3)
- More diagnostic and formative testing should be used. (2)
- Competency-based assessment should be introduced.* (2)
- Some schools' assessment programs are inadequate.* (2)

A detailed summary of all responses to this question is presented in Appendix 4.

---

* Items marked with an asterisk were mentioned by both inspectors and respondents in schools.

IV SUMMARY

The *major aims* of the study were to systematically map the present practices relating to assessment in a sample of schools and to determine the degree of satisfaction among respondents in these schools with these practices. Where appropriate, alternative procedures desired by respondents in schools were explored.

Clarification of objectives within the school

Respondents were asked about the clarification of objectives within the school. Objectives which related to the school as a whole were distinguished from those which related to specific subjects within the school.

All schools had definite policies related to some aspects of assessment. This usually took the form of maximum limits on the frequency and/or length of assessments.

All schools are required to submit to the Board a form (Ml) for each Board subject stating the assessment program for that subject in that school. Almost all teachers indicated that the Ml form was readily accessible to them. The majority of teachers based their teaching program on the workbook, although half of the Maths I teachers used text books as the main basis for their teaching.

Summative assessment

In this study, the summative assessment program is defined to be all of the tests, assignments and practical work which contribute to the exit measures of student achievement produced by the school. Information was obtained concerning the structure of the summative assessment program and changes desired by various respondents. The decision making processes were also investigated. The topics covered included:

- time allowed for assessments
- weighting on various assessment elements
- weighting on modes of assessment used
- weighting on types of items used
- use of common tests (within and between schools)
- setting of assessment instruments.

Time allowed

The time devoted to assessment was similar between subjects within schools but varied markedly between schools.

Weighting on various assessment elements, modes of assessment and types of items

The weighting given to various assessment elements, modes of assessment and types of items varied considerably both between subjects and between schools. However, teachers were largely responsible for the choice of what was used, and consequently, were generally satisfied in these areas.

In Maths I, the ratings were based almost entirely on formal tests under examination conditions. Extended response items (mathematics problems) were used almost exclusively. All schools in the sample had an end of semester test with a weighting of at least 40 per cent. All but one of these schools had a mid-semester test which accounted for from 20 to 40 per cent of the weighting. In most schools, there were no other elements in the summative assessment program in Maths I. The major exception was the computer maths unit where assignments were part of the assessment program. Inspectors' opinions on the desirable features of a summative assessment program in mathematics corresponded closely with present practices.

In English, the ratings were based on a large number of assessments of a widely varying nature in terms of mode and types of items. All schools visited had at least one test under examination conditions at the end of the semester, and at least one such test during the semester. In all schools, there was an oral/aural component and a component for unsupervised work completed at home. Generally, assessment in English occurred more frequently during the semester than in Maths I.

In foreign languages, the ratings were also based on a large number of assessments of a widely varying nature. The assessment structure was similar to that for English with a greater emphasis on oral/aural assessment. In general, assessment occurred more frequently in the foreign languages than in either English or Maths I.

The frequency of assessments suggested by inspectors in all three subjects broadly corresponded with assessment practices in these subjects.

Whereas the majority of principals and teachers were satisfied with the present weightings given to elements of assessment, modes of assessment and types of items, about one-third were dissatisfied with each area. However, there appeared to be no consistent trend in the dissatisfactions expressed.

In general, a large proportion of the items were 'teacher-made', particularly the essays and assignments. The only types of items which were extensively 'borrowed' were the multiple choice and completion/short answer items, and the extended response problems in Maths I.

Use of common tests

Within all schools, common maths tests were used where two or more classes were studying the same unit. Mathematics teachers appeared satisfied with this arrangement. A large proportion of these teachers indicated that common tests between schools would be desirable to achieve comparability between schools. While some envisaged this as a form of reference test, others desired a return to an external examination.

While the percentage of assessment that was common across Year 12 English classes within a school varied markedly, most schools used more than 70 per cent common assessment between classes in English. Inspectors felt that in English there should be some common and some different tests between classes. Almost all English teachers indicated that they did not want common testing between schools. In general, these teachers believed that they could achieve comparability between schools through discussion and comparison of marking standards. This method was also used within schools to obtain comparability in English where different tests were used between classes. In most schools, each English teacher marked part of the common assessment across all classes. In addition, some schools used a team-teaching approach so that each teacher taught a particular section (e.g. drama) to all of the classes.

No school had more than one class doing the same foreign language at Year 12 level. Some foreign language teachers suggested common tests between schools in areas of the course such as grammar and comprehension. However, none of these teachers suggested a return to full external assessment.

About half of the principals allowed some different testing between classes in some subjects. In schools which had all common tests between classes, the principals expressed no desire to introduce different tests. Half the principals opposed any form of common testing between schools, as they felt that the present flexibility available to schools would be lost.

Most mathematics teachers, very few English teachers and about half of the foreign language teachers indicated that some form of common testing between schools would be desirable. Only half of the principals supported any form of common testing between schools. The only support for a return to an external Senior-type examination came from the mathematics teachers.

Setting of assessment instruments

In all cases, the school's subject area co-ordinator for each subject decided who prepared each assessment instrument in that subject. A wide range of criteria were used in determining this. In almost all cases, respondents indicated that they were satisfied with the allocation of who sets the tests.

Overview of summative assessment

Some patterns showing both similarities and differences among subjects emerged in the extent and nature of the summative assessment programs used in schools.

In general, teachers were satisfied with the present summative assessment program for their Year 12 classes. Those who did desire change indicated that they wanted a small change in emphasis rather change in its structure. In English, there was some support for an increased emphasis on oral work and other communication skills with a decreased emphasis on literature. However, many of those who wanted changes indicated that the changes were currently being achieved in Year 11 through the implementation of the new syllabus.

Most decisions concerning the structure of the summative assessment program were made by the subject masters, usually in consultation with the other teachers. In all cases, the principal could over-rule the decisions made by subject masters, but it appeared that this power was seldom exercised. The pattern of decision making thought desirable by inspectors corresponded substantially with what appears to be occurring in schools.

Formative assessment

In this study, formative assessment includes any assessment which provides feedback to teachers and students regarding student progress. Two of the aspects of formative assessment probed in the study were:

- the use of summative assessment for formative purposes
- the balance between summative and formative assessment.

The use of summative assessment for formative purposes

Whereas the major purpose of the elements in the summative assessment program is the production of exit measures of achievement, these assessments also provide some feedback on student performance. That is, there is a formative aspect to summative assessment. The effectiveness of the feedback derived from the summative assessment program is affected by various factors. Some of these factors were investigated in this study. The longer the delays between submitting work and receiving marks and comments, the less effective is the feedback. When students are interested in where and why they lost marks, feedback is more effective. Similarly, where students are encouraged to make corrections on their papers and take them home, the feedback is more effective.

Teachers and students generally agreed that the delay in getting Maths I and foreign language papers back as satisfactory. In general, for both subjects, papers were returned within one week. In English, the delays were longer than in the other two subjects. These delays in English were considered unsatisfactory by students in about half of the schools. In contrast, almost all English teachers perceived the delay as being satisfactory.

Only a few students indicated that they were interested in their score rather than where and why they lost marks. However, teachers felt that students were more interested in their score. The principals were equally divided in their opinions on this issue.

In some cases, students were allowed to make corrections in their papers while in other cases the papers were kept 'clean' for moderation meetings. This seemed to depend on the individual teacher rather than the school or the subject. Some schools had definite policies that papers must be taken home and shown to parents, while most schools discouraged the taking home of papers.

The balance between summative and formative assessment

The majority of inspectors, principals and teachers indicated that they were satisfied with the present frequency of summative assessments. Overall, about one-third of each level of respondent desired to decrease the amount of summative assessment. While students are reasonably satisfied with the summative assessment load in these three subjects, some informal comments by students indicated a heavy workload for those students doing mainly humanities subjects, although data relating to these subjects was not sought in this study.

Principals interviewed in this study appear acutely aware of the problems of over-testing and many are taking steps to ensure that over-testing does not occur.

The frequency of summative assessments in English and foreign languages was considerably greater than for Maths I. In Maths I, students tended to want tests to be smaller and more frequent to provide more frequent feedback on their progress, whereas English and foreign language students tended to request a slightly decreased assessment load. When asked if they would like to change the emphasis on summative and formative assessment, many teachers indicated a preference for a small shift from summative to formative, but generally they did not see the summative assessment program as dominating their teaching. Many indicated that they were happy with the present emphasis placed on summative and formative assessment. Half of the principals and most of the inspectors wanted to increase the amount of formative assessment. Most teachers in the subjects studied indicated that they were able to integrate formative assessment into their teaching programs.

Marking and validation procedures

Allocation of marking

The method of allocation of marking varied within schools from subject to subject. In most schools, the class teacher marked the Maths I papers of his own class, while in English, each teacher does part of the marking across all classes.

Use of marking schemes

Almost all Maths I and foreign language teachers used detailed, prepared marking schemes for all assessments. Most English teachers had a detailed marking scheme for such things as comprehension while essay marks were usually assigned according to a global perception of the standard of the work.

Comparability between classes

Students in a number of schools felt that comparability between classes was not achieved in English, but they did not perceive any lack of comparability between classes studying the same unit of Maths I in the same semester within a school. Most principals and teachers did not consider the achievement of comparability between classes to present any problems. This opinion was generally shared by the inspectors. However, this confidence is not shared by writers in the field of educational measurement.

Consistency of marking from year to year

Almost all teachers believed that their marking standards were consistent from year to year. They believed that this consistency was achieved through a combination of their subjective judgement and the re-use of some test items. However, this re-use of items did not appear to be based on any educational measurement criteria.

Validity and reliability of assessment instruments

Most principals and teachers indicated that they were satisfied with the quality of the assessment instruments in their school or subject. However, on being questioned further, most indicated that their view was based purely on their own perceptions rather than on any specific validation procedure. In contrast to the apparent satisfaction of respondents in schools with the quality of tests, only one of the inspectors interviewed was satisfied with their quality.

When asked whether they felt that teachers possessed the necessary skills to ensure that testing was valid and reliable, two inspectors said that teachers did not, and the other four inspectors felt only some teachers possessed these skills. Research studies into testing techniques support these views held by the inspectors.

Achievement and other exit measures

Data was collected on the following aspects of measuring and reporting student achievement:

- moderated school assessments
- special subject assessments
- school reports

Moderated school assessments

Derivation of final scores:

In the series of booklets prepared by the Department of Education, to guide teachers' developing strategies to assist in implementing the moderation system, the point was made that to achieve the desired weightings it is necessary to allow for differences in standard deviation between tests when combining test scores. In this study, respondents in schools indicated that final scores used to determine ratings were obtained by merely adding raw scores. However inspectors were aware of the effect of differences in standard deviations and all but one inspector recommended that standardised scores be added rather than raw scores.

The moderation system:

About two-thirds of the respondents were dissatisfied with the moderation system. Many respondents in English wanted an alternative which decreased the emphasis placed on moderation, giving schools more responsibility for maintaining standards. Very few of these respondents explained any ramifications of their proposed change. Half of the 20 Maths I respondents wanted a return to external examinations. Some of these wanted internal assessment as well. Most of those who proposed a return to external examinations had a reasonable perception of the positive effects of such a move, although only one mentioned any of the negative effects. About half of foreign language teachers wanted either a reference test or external assessment of some parts of the course, such as grammar. These teachers seemed to be aware of both the positive and negative effects of their proposals.

Most of the movements of ratings recommended by the moderation meetings were downward movements. However, only about half of the changes recommended were subsequently made.

Determination of student ratings:

For each of the groups of respondents in schools, the major criterion used in determining student ratings was teacher perception of standards. Inspectors also felt that teacher perception of standards should be the main criterion used to determine cut-off marks.

Agreement at moderation meetings requires that the reviewers have similar perception of standards to staff in the school submitting the papers. However, research does not support the notion that teachers are able to judge standards of student achievement. It is possible that the consensus reached at moderation meetings stems from the need to make a decision rather than from a belief that comparability has been achieved.

Special Subject Assessments (SSAs)

Derivation of SSAs:

There appeared to be great diversity in the procedures used by schools to assign the SSA which is sent to the Board for input into the TE Score calculations. Only half of the 12 schools used standardisation effectively when combining semester scores in calculating the SSA.

It appeared also that very few teachers have much knowledge of, or interest in, the procedures used for these calculations in their school. Usually major responsibility for the calculation is given to one or two persons (generally, the principal, deputy principals, the mathematics subject master or a mathematics teacher) with a small degree of participation required from the other subject masters. While principals reported that input from 'other teachers' occurred frequently, these teachers felt that an input from them was seldom sought.

Manipulation of SSAs:

Several interviewees admitted that they used procedures to attempt to manipulate the SSAs of potential tertiary entrants. Most respondents were of the opinion that such manipulations were possible. Quite a few of these felt it was being done by some schools but did not believe it was being done in their school. However, some respondents indicated that they had manipulated SSAs in their school.

Fairness of TE Scores:

Although about half of the 70 respondents said that some groups were unfairly affected in calculation of TE Scores, few of them were able to justify their answers in terms of the procedures used in calculating the TE Scores. The responses isolated three areas of concern: the nature of the subject; the ability level of students studying the subject; and the size of the group studying the subject.

In view of this apparent ignorance of and lack of confidence in the TE Score, it appears that further research to examine some of the pertinent issues should be carried out.

School Reports

The content of the end of semester report varied markedly among schools. All schools provided at least one achievement measure, and all except one school provided at least one affective measure. While most schools appeared to use coded comments, most inspectors felt the use of written comments was much more desirable. Several schools also reported to parents after mid-semester tests. This report was either written or oral, and was usually less formal than that given at the end of the semester. Class teachers were responsible for assigning almost all descriptors on the reports.

Although about half of the respondents in schools were dissatisfied with the information on the school report, there did not appear to be a pattern among the suggested changes.

General concerns and suggestions

Respondents were given the opportunity to elaborate on issues discussed earlier in the interview or to discuss issues not dealt with previously. The most frequently mentioned general concerns involved over-testing, lack of comparability between schools, misrepresentation of student work by teachers at moderation meetings and a desire to return to external examinations. While over-testing did not appear to be a problem in the individual subjects studied, the cumulative effect of the different assessment programs across the school was a cause of concern. This problem appeared to effect only a minority of students, namely those studying courses involving mostly humanities subjects.

Further aims of the study were to examine the effects of the requirements of outside bodies on the assessment programs within the school and to compare the assessment situation in 1978 with that in 1974 where possible.

Since many of the comparisons between 1974 and 1978 involve the effects of outside bodies on the school, the findings for these two aims are discussed together in this summary.

Formative and summative assessment

Data collected in 1974 in the Schools Under Radford Study indicated that most teachers held very strong views about the emphasis on summative assessment which, they felt, prevented them from devoting any time to formative assessment. This contrasts sharply with the present study which indicates that, in 1978, formative assessment was frequently used, and that over-testing is no longer a serious problem in the subjects investigated.

In the Schools Under Radford Study, teachers commented that students' marked scripts were usually retained by the school for presentation at moderation meetings. Thus, they perceived this constraint as severely limiting the use of summative assessments for formative purposes. In 1978, school staff and students generally agreed that test papers were given back to students after tests in the three subjects investigated. As a result, these teachers felt that they were able to provide useful feedback to students from the summative assessments.

Flexibility in assessment procedures

The results of the Schools Under Radford Study suggested that teachers in 1974 were not taking advantage of the flexibility in assessment procedures afforded by the implementation of the Radford Report in 1971. By contrast, in 1978, some teachers, particularly those of English and foreign languages, appear to have capitalised on the flexibility available through internal assessment with regard to both modes of assessment and types of item. However, moderation meetings still base their judgements mainly on written work. This emphasis on written work still appears to be constraining teachers to limit their use of oral assessment techniques.

Flexibility in achievement measures

A considerable number of respondents have doubts about the achievement of comparability between schools. Many teachers felt that once ratings had been given in semester one, it was very difficult to adjust the distribution upward in subsequent semesters if some students showed a marked improvement in their achievement.

As mentioned above, there was considerable diversity in the procedures used in assigning SSAs. This reflects the Board's attitude that schools should be responsible for determining the method of deriving SSAs. That is, schools are not constrained to use a certain method, nor do they feel constrained to do so.

The researchers did not perceive any constraints on the information contained in school reports as a result of pressures from employers.

The *last aim* of the study was to determine the extent to which school staff believe that the present system facilitates the achievement of the purposes of assessment.

To determine this, principals and teachers were asked to focus upon the five assessment objectives listed in the *Review of School-Based Assessment in Queensland Secondary Schools*. These objectives are:

- to enable the student to assess his/her learning and to provide information which can be used to correct deficiencies in his/her learning;
- to enable the teachers and students to modify teaching/learning programs to suit the needs of the students;
- ultimately to provide the students and parents with information concerning the students' achievement and potential;
- to assist students and parents in determining future education and employment pathways; and
- to provide, for other educational institutions and employers, an indication of the suitability and readiness of the student to undertake further education and/or employment.

In this study, principals and teachers were asked to indicate the extent to which present assessment procedures and exit measures facilitated, militated against, or had no effect on, the achievement of these objectives. The majority of these respondents perceived these aspects to be either facilitated or having no effect on the achievement of these stated objectives. Only a few respondents perceived aspects of the current assessment procedures as militating against the achievement of these objectives.

This discussion about the purposes of assessment from the *Review of School-Based Assessment in Queensland Secondary Schools* report frequently led to incidental comment about that report. These discussions indicated that very few interviewees had read either the report or the summary of its recommendations that has been distributed to schools. Therefore, teachers largely seem to have little knowledge of the implications for its implementation. Of those who are reasonably familiar with its recommendations, about half like the idea of competency-based rather than norm-based assessment. However, most feel the comparability problems will be just as great as at present, and feel that in addition, the workload on teachers will be greatly increased.

On the basis of the above discussions in schools, it appeared neither teachers nor administrators feel that the implementation of the recommendations of the report *Review of School-Based Assessment in Queensland Secondary Schools* would alleviate the present problems associated with the production of exit measures of achievement at the Year 12 level in schools. Many felt that replacing the present moderation system by a system of accreditation and using a verbal system of reporting will increase confusion and reduce public acceptance of the Queensland education system.

V CONCLUSIONS

There were two major aspects to this report - firstly, a detailed examination of present assessment practices in specific subject areas and secondly, an interpretation of the trends in, and effectiveness of, the present assessment system operating at the Year 12 level in Queensland State secondary schools.

The major findings relating to the first aspect are:

- Most decisions relating to assessment procedures in the subjects studied are made by the subject master or subject area co-ordinator in consultation with other teachers of the subject. The principal has the power to veto decisions, but it appears that he seldom exercises it.
- Teachers are generally satisfied that mechanical aspects of assessment and moderation are functioning fairly smoothly.
- There is substantial concern that comparability between schools is not being achieved.
- Schools are utilising the flexibility available to them in their choice of modes of assessment.
- Although outside bodies (particularly the Board) make substantial demands on schools, respondents generally felt that summative assessment did not dominate their assessment program and that they were generally able to integrate formative assessment into their teaching program.
- There is considerable diversity in the procedures used by schools to assign the Special Subject Assessment (SSA) which is sent to the Board for input into the Tertiary Entrance (TE) Score calculations. Only half of the schools used standardisation effectively when combining semester scores in calculating the SSA.
- Most principals and teachers indicated that they were satisfied with the quality of the assessment instruments in their school or subject. However, most indicated that their view was based purely on their own perceptions rather than on any specific validation procedure. In contrast to the apparent satisfaction of respondents in schools with the quality of tests, only one of the inspectors interviewed was satisfied with their quality. Inspectors believed that many teachers did not possess the skills necessary to ensure that tests were valid and reliable.

With regard to the second aspect of this study, the researchers believe that:

- Between 1974 and 1978, there have been considerable improvements in:
    - the use of formative assessment
    - the use of appropriate modes of assessment
    - the mechanics of the moderation system
- Overall, there is greater satisfaction with the system at present than there was in the years immediately following the implementation of internal assessment.
- Further research into the method of calculation of TE Scores should be undertaken.
- Further research into the validity and reliability of assessment programs and instruments used in schools should be carried out.

APPENDIX I

Mapping of Assessment Procedures in Schools

INTERVIEW SCHEDULE - INSPECTORS

BACKGROUND INFORMATION:

1. Name:
2. Subject:
3. Are you on a subject advisory committee?

YES	
NO	

(The numbering in this schedule corresponds with the numbering of parallel questions in the Principal and Teachers Interview Schedule.)

① CLARIFICATION OF OBJECTIVES WITHIN THE SCHOOL

B. COURSE OBJECTIVES

4. Should teachers teach from:

Syllabus	
M1	
Workbook	
Textbook	
Other	

If other, what?

5. Should assessment be based on:

Teaching program	
Syllabus	
M1	
Workbook	
Textbook	
Other	

If other, what?

6. Are there any course objectives which cannot adequately be assessed?

YES	
NO	

If YES, what?

② SUMMATIVE ASSESSMENT

A. STRUCTURE OF SUMMATIVE ASSESSMENT

What is most desirable? What is an acceptable range of assessment?

	TEST NUMBER										
1.	When administered? (Which week of semester?)										
2.	Time allowed? (mins.)										
3.	Weighting of elements of assessment? (%)										
4.	Test common across classes or not? (YES/NO)										
5. Mode of Assessment	Exam conditions										
	Supervised collaboration (in class)										
	Unsupervised (at home)										
	ORAL										
	WRITTEN										
	OTHER										

Type of item?		(Tick=Yes)	Approx. weight (%)
Multiple-choice			
Completion (short answer)			
True-false			
Extended response (Essay/maths problem)			
Assignment	In class		
	At home		
	Group		
Practical work			
	Debate		
	Drama		
	Conversation		
	Listening Comprehension		
	Dictation		
	Reading a Passage		
	Other		

B. DECISION MAKING FOR SUMMATIVE ASSESSMENT

		Who should decide?	What information should be sought during decision making?
1.	When administered		
2.	Time allowed		
3.	Weighting of elements of assessment		
4.	Common/ different tests (within school)		
5.	Mode of assessment		
6.	Types of items		

D. THE SETTING OF ASSESSMENT INSTRUMENTS

1. Who should set the assessment instruments?

SM	
Teacher	

3. On what basis should the decision be made?

Teaching experience	
Rotational	
Each set one section	
Content expertize	
Time constraints	
Other	

If other, what?

5. Are you happy with the quality of assessment instruments in your subject?

YES	
NO	

6. What steps should be taken to ensure that tests are valid and reliable?

7. Do teachers possess the necessary skills to take these steps?

YES	
NO	

(3.) FORMATIVE ASSESSMENT

A. FORMATIVE ASPECTS OF SUMMATIVE ASSESSMENT

1. Should students get their test papers back?

YES	
NO	

If yes, what is the maximum delay that would be satisfactory?

3. Should papers be worked through:

. in class?

YES	
NO	

4. Should students be able to:

. make corrections on their papers?

YES	
NO	

. take their papers home?

YES	
NO	

B. FORMATIVE vs. SUMMATIVE ASSESSMENT

1. Should there by any assessment which does not count towards ratings?

YES	
NO	

If yes, approximately how much?

2. Should the present emphasis on each type change?

	incr.	same	decr.
Summative (count toward ratings)			
Formative (do NOT count)			

If change is desired,

. for what reasons?

. what would be the effects?

THOSE AFFECTED	+ EFFECTS	- EFFECTS
students		
parents		
teachers		
school administration		
employers		
tertiary institutions		
Board of Secondary School Studies		
Other		

---

④ MARKING PROCEDURES

1. Who should mark the papers?

Teacher marks own class	
Each teacher marks a section	
One teacher marks all	
Other	

If other, who?

2. Should there be a detailed, prepared marking scheme?

YES	
NO	

3. Should the marking of one teacher be perused by another?

YES	
NO	

4. Should any other steps be taken to ensure comparability between classes?

YES	
NO	

If yes, what steps?

5. Do you feel that comparability is achieved between:

.classes within schools?

YES	
NO	

.schools?

YES	
NO	

---

⑤ ACHIEVEMENT AND OTHER EXIT MEASURES

A. MODERATED SCHOOL ASSESSMENT

1. How should test scores be combined to obtain the score for the semester?

add (weighted) raw scores	
add (weighted) standardized scores	

If other, what?

4. What criteria should individual schools use for the cut-off and distributions of ratings on the M4?

normal curve	
cut-off at fixed percentages (total score)	
teacher perception of standards	
distribution of these students last year	
distribution of last year's Grade 12	
other	

8. Are you happy with the present procedures?

If not: what alternatives would you suggest?

YES	
NO	

. for what reasons?

. have you suggested changes?

If yes, to whom and with what result?

YES	
NO	

- what would be the effects?

THOSE AFFECTED	+ EFFECTS	- EFFECTS
students		
parents		
teachers		
school administration		
employers		
tertiary institutions		
Board of Secondary School Studies		
Other		

B. SPECIAL SCHOOL ASSESSMENT

3. How should the S.S.A. be prepared?

Semester	1	2	3	4
Scores input				
Weighting				

- After calculation, are the scores given back to the class teachers/subject master for perusal and/or modification?

YES	
NO	

4. Do you think it is possible to manipulate the S.S.A. to advantage some students?

YES	
NO	

If yes, . how can it be done?

. to what extent do you think it is done?

6. Which groups, if any, do you feel are unfairly effected by the assigning of S.S.A.'s and subsequent calculation of T.E. scores?

Group	Adv. (A) Disadv. (D)	Why?

C. SCHOOL REPORT

1. What information should go on the school report?

⑦ GENERAL CONCERNS AND SUGGESTIONS

1. What aspect of assessment, if any, causes you considerable concern?

2. Why?

3. What do you think could be done about it?

APPENDIX 2

Mapping of Assessment Procedures in Schools

INTERVIEW SCHEDULE - PRINCIPALS and TEACHERS

BACKGROUND INFORMATION:

1. Name of School:

2. Interviewee:

Principal	
Deputy Principal	
Subject Master	
Other Teacher	

Subject Area Co-ordinator	
Other Teacher	

3. Teaching Experience:

1st Year	
2nd Year	
3rd-5th Year	
6th or later Year	

4. Subject:

① CLARIFICATION OF OBJECTIVES WITHIN THE SCHOOL

A. SCHOOL POLICIES related to assessment:

1. Are there any?

YES	
NO	

If yes, . what are they?

. how are these communicated to you?

B. COURSE OBJECTIVES:

1. Who prepares the M1?

ADMIN.	
S.A.C.	
O.T.	

2. How frequently do you need to refer to the M1 during the semester?

3. Is it accessible to you when you wish to refer to it?

YES	
NO	

If no, comment.

4. Do you teach from:

Syllabus	
M1	
Workbook	
Textbook	
Other	

If other, what?

5. Is the assessment based on:

Teaching program	
Syllabus	
M1	
Workbook	
Textbook	
Other	

If other, what?

6. Are there any course objectives which cannot adequately be assessed?

YES	
NO	

If yes, what?

Type of item?		Used? (Tick= Yes)	Approx. weight? (%)	Source? Teacher-made (%)	"Borrowed" (%)
Multiple-choice					
Completion (short answer)					
True-false					
Extended response (Essay/maths problem)					
Assign-ment	In class				
	At home				
	Group				
Practical work					
	Debate				
	Drama				
	Conversation				
	Listening Comprehension				
	Dictation				
	Reading a Passage				
	Other				

---

② SUMMATIVE ASSESSMENT

A. STRUCTURE OF SUMMATIVE ASSESSMENT

TEST NUMBER										
1. When administered? (Which week of semester?)										
2. Time allowed? (mins.)										
3. Weighting of elements of assessment? (%)										
4. Test common across classes or not? (YES/NO)										
5. Mode of Assessment?	Exam conditions									
	Supervised collaboration (in class)									
	Unsupervised (at home)									
	ORAL									
	WRITTEN									
	OTHER									

B. DECISION MAKING FOR SUMMATIVE ASSESSMENT

		Who decides?	What information is sought by you/from you during decision making?	How are decisions communicated back by you/to you?
1.	When administered			
2.	Time allowed			
3.	Weighting of elements of assessment			
4.	Common/ different tests (within school)			
5.	Mode of assessment			
6.	Types of items			

C. PROPOSALS FOR CHANGE

Would you like changes made in:

			YES/NO	What changes?	Why?
1.	When administered				
2.	Time allowed				
3.	Weighting of elements of assessment				
4.	Common/ different tests	within schools			
		between schools			
5.	Mode of assessment				
6.	Types of items				

7. Have you made any of these recommendations to anyone?

YES	
NO	

If yes, . to whom?

. what would be the effects?

THOSE AFFECTED	+ EFFECTS	- EFFECTS
students		
parents		
teachers		
school administration		
employers		
tertiary institutions		
Board of Secondary School Studies		
Other		

D. THE SETTING OF ASSESSMENT INSTRUMENTS

1. Who sets the assessment instruments?

INTERVIEWEE	
OTHER	

ADMIN	
SM	
TEACHER	

2. Who decides who sets the tests?

3. On what basis is the decision made?

Teaching experience	
Rotational	
Each set one section	
Content expertize	
Time constraints	
Other	

If other, what?

4. Would you like to change this allocation of who sets the tests?

If yes, . why?

YES ☐
NO ☐

. what system of allocation would you prefer?

5. Are you happy with the quality of assessment instruments at this school?

YES ☐
NO ☐

6. What steps are taken to ensure that tests are valid and reliable?

(3.) FORMATIVE ASSESSMENT

A. FORMATIVE ASPECTS OF SUMMATIVE ASSESSMENT

1. Do students get their test papers back?

YES ☐
NO ☐

If yes, what is the delay and is it satisfactory?

. marks SATISFACTORY/UNSATISFACTORY
. scripts SATISFACTORY/UNSATISFACTORY

2. What proportion of students are interested in:

. marks only?
. reasons for marks given/marks lost?

3. Are papers worked through:

. in class?

YES ☐
NO ☐

. individually with students?

YES ☐
NO ☐

4. Can students:

- make corrections on their papers?

YES	
NO	

- take their papers home?

YES	
NO	

B. FORMATIVE vs. SUMMATIVE ASSESSMENT

1. Is there any assessment which does not count towards ratings?
   If yes, approximately how much?

YES	
NO	

2. Should the emphasis on each type change?

	incr.	same	decr.
Summative (count toward ratings)			
Formative (do NOT count)			

If change is desired,

- for what reasons?

- how could this be achieved?

- what would be the effects?

THOSE AFFECTED	+ EFFECTS	- EFFECTS
students		
parents		
teachers		
school administration		
employers		
tertiary institutions		
Board of Secondary School Studies		
Other		

④ MARKING PROCEDURES

1. Who marks the papers?

Teacher marks own class	
Each teacher marks a section	
One teacher marks all	
Other	

If other, who?

2. Is there a detailed, prepared marking scheme?

YES	
NO	

3. Is the marking of one teacher perused by another?

YES	
NO	

4. Are any other steps taken to ensure comparability between classes?

YES	
NO	

If yes, what steps?

5. Do you feel that comparability between classes is achieved?

YES	
NO	

6. Is your marking standard consistent from year to year?

YES	
NO	

If yes, how do you achieve this?

7. Who decides on marking procedures?

8. Are you happy with these marking procedures?

YES	
NO	

If not:

. in what areas?

. for what reasons?

. have you proposed alternatives?
If yes: to whom?

YES	☐
NO	☐

: with what success?

(5.) ACHIEVEMENT AND OTHER EXIT MEASURES

A. MODERATED SCHOOL ASSESSMENT

1. How do you combine test scores to obtain the score for the semester?

add (weighted) raw scores	☐
add (weighted) standardized scores	☐
other	☐

If other, what

2. Who decides on the method of combination?

3. Who decides on the cut-off marks on the M4?

4. What criteria are used for the cut-off and distributions of ratings on the M4?

normal curve	☐
cut-off at fixed percentages (total score)	☐
teacher perception of standards	☐
distribution of these students last year	☐
distribution of last year's Grade 12	☐
other	☐

If other, what?

5. What was the nature of comments given at the moderation meeting?

Critical	
Constructive	
Irrelevant	
Useful	
Other	

If other, what?

6. Did you feel the comments were justified?

YES	
NO	

Why?

7. Did the moderation meeting recommend a change in the distribution of ratings?

YES	
NO	

If yes: what was done to obtain the final distribution?

. what criteria were used?

. who made the final decision?

8. Are you satisfied with the present moderation procedures?

YES	
NO	

If not: what alternatives would you suggest?

. for what reasons?

. have you suggested changes?
If yes, to whom and with
what result?

YES	
NO	

. what would be the effects?

Those affected	+ Effects	- Effects
students		
parents		
teachers		
school administration		
employers		
tertiary institutions		
Board of Secondary School Studies		
Other		

B. SPECIAL SCHOOL ASSESSMENT

1. Who prepares the S.S.A. in your subject?

2. Is there a school policy about:

. who prepares the S.S.A.?

YES	
NO	

. how the S.S.A. is prepared

YES	
NO	

3. How is the S.S.A. prepared?

Semester	1	2	3	4
. Scores input				
. Weighting				

. Standardized?

	Before Input	After input
YES		
NO		

. Do you use last year's A.S.A.T. means to estimate T.E. scores for this year's group?

YES	
NO	

. After calculation, are the scores given back to the class teachers/subject master for perusal?

YES	
NO	

. Other comment?

4. Do you think it is possible to manipulate the S.S.A. to advantage some students?

YES	
NO	

If yes, . how can it be done?

. to what extent do you think it is done?

5. Which of the following do you take into account in calculating the S.S.A.:

	YES/NO	Why? In what way?
mean		
standard deviation		
shape of the distribution		

Other comments:

6. Which groups if any, do you feel are unfairly effected by the assigning of S.S.A.s and subsequent calculation of T.E. scores?

Group	Adv. (A) Disadv. (D)	Why?

C. SCHOOL REPORT

1. What information goes on the school report?

2. Who assigns the marks/ratings reported?

3. How are the marks/ratings calculated or assigned?

4. Are you satisfied with the information on the school report?

YES	
NO	

If no, what changes would you propose?

. why?

. what would be the effects?

Those affected	+ Effects	- Effects
students		
parents		
teachers		
school administration		
employers		
tertiary institutions		
Board of Secondary School Studies		
Other		

5. Have you suggested any changes?

YES	
NO	

If yes, to whom?

---

⑥ ATTAINMENT OF THE PURPOSES OF ASSESSMENT

Do the following facilitate (F.), militate against (M.A.), or have no effect (N.E.) on the purposes of assessment identified by the Scott Committee.

ASSESSMENT OBJECTIVES	TESTING PROGRAM			MODERATED SCHOOL ASSESSMENT			T.E. SCORES			SCHOOL REPORTS		
	F.	N.E.	M.A.	F.	N.E.	M.A.	F.	N.E.	M.A.	F.	N.E.	M.A.
1. assess learning and correct deficiencies												
2. to modify learning to meet student needs												
3. to measure student achievement and potential												
4. to assist students and parents determine future pathways												
5. to provide a measure of suitability for education and employment												

(7.) GENERAL CONCERNS AND SUGGESTIONS

1. What aspect of assessment, if any, causes you considerable concern?

2. Why?

3. What do you think could be done about it?

APPENDIX 3

Mapping of Assessment Procedures in Schools

INTERVIEW SCHEDULE - STUDENTS

BACKGROUND INFORMATION:

1. School:

3. Class:

3. Student numbers per subject:

Subject	Number	Teacher	Assign-ments YES/NO
English (E)			
Mathematics I (M)			
Foreign Languages (F)			

			Tests			Assignments		
			E	M	F	E	M	F
1.	Do you get your test papers/ assignments back?	YES						
		NO						
	If YES, how long (in days) before you receive them?	1-2						
		3-7						
		8-14						
		> 14						
2.	Overall, do you feel that the time between submitting tests/assignments and receiving the results is:	satis-factory?						
		too long?						
3.	Do teachers work through tests: . in class?	YES						
		NO						
	. individually?	YES						
		NO						
4.	When test papers are returned can you: . make corrections on them?	YES						
		NO						
	. take them home?	YES						
		NO						
	If NO, would you like to take them home?	YES						
		NO						
5.	Are the numbers of tests/ assignments which contribute to ratings:	too many?						
		adequate?						
		too few?						
6.	Do you have any tests/ assignments which do NOT count toward ratings?	YES						
		NO						

If YES: . how seriously do you take them?

		E	M	F
. do you find them useful?	YES			
	NO			

			Both tests and assignments		
			E	M	F
7.	Would you like a change in the number of tests and assignments which do not count towards ratings, but which help you understand how well you are doing and tell you what sort of mistakes you are making?	YES, MORE			
		NO			
		YES, FEWER			
8.	When you get tests/assignments back, are you interested in marks only, or in finding out where and why you lost marks?	marks only?			
		why marks were lost?			
9.	Who marks your tests/assignments? . you class teacher				
	each section is marked by a different teacher				
	. all papers are marked by the one teacher				
	. other (specify)				
10.	Are you happy about who does does the marking?	YES			
		NO			
	If NO, what would you prefer?				
11.	Are you given a detailed marking scheme?	YES			
		NO			
12.	Would you like a more detailed marking scheme?	YES			
		NO			
13.	Do you feel that the end of semester ratings are comparable: . between classes?	YES			
		NO			
	. between schools?	YES			
		NO			

COMMENTS:

APPENDIX 4

ADDITIONAL AREAS OF CONCERN AND SUGGESTED CHANGES

This data summarises the responses to the final question. Respondents were given the opportunity to elaborate on issues discussed earlier, if they desired, or to discuss concerns not dealt with previously. Thus, there is some repetition of data presented elsewhere.

The responses have been classified into four categories. The numbers in brackets indicate how many respondents mentioned each topic in response to this question. The comments of inspectors are listed in *italics* to distinguish them from the comments of principals and teachers because the function of inspectors in this study differs from that of principals and teachers. Comments made by both groups have been marked with an asterisk (*).

1. Problems with the Present System

There is too much formal testing and this is causing undue pressure on students. (15)*

Comparability among schools is not being achieved. (10)

Some teachers misrepresent the levels of achievement of their students, in order to obtain higher ratings. (10)

The quality of tests varies considerably among schools, with the assessment programs at some schools being inadequate. (6)*

Moderation encourages schools to teach and test only materials which is easily moderated. (4)

Moderation meetings waste too much teaching time. (3)*

An expert review panel which visits schools should replace moderation meetings. (2)

Norm-based assessment emphasises competition rather than the actual standard of work. (3)

Teachers are uncomfortable about being responsible for determining their students' futures. (2)

The delays in receiving Board approved ratings are too long. (2)

Present moderation procedures do not allow enough time to moderate effectively. (2)

It is sometimes difficult to store work for, and transport work to moderation meetings. (1)

Schools cannot serve two masters simultaneously (i.e. the Board of Secondary School Studies and the Department of Education). (1)

There is not enough emphasis on formative assessment. (1)*

*There is too much formal testing. (3)**

*More diagnostic and formative testing should be used. (2)**

*Some schools assessment programs are inadequate. (2)**

*The mechanics of moderation waste time. (1)**

The information on the M4 often contradicts the plan given on the M1. (1)

2. Proposed Alternative Systems

External examinations would be fairer. (10)

The quality of work is lower now than it was under external exams. (1)

Competency-based assessment is desirable. (5)*

Competency-based assessment would increase the workload of teachers and administrators. (4)

Accreditation of schools would overcome most moderation problems. (2)*

*Competency-based assessment should be introduced. (2)**

Competency-based assessment will not in itself improve the situation. (1)

*Accreditation of schools is desirable. (1)**

Accreditation of schools is impractical and elitist. (1)

3. Subject Specific Problems

Oral work is difficult to assess. (5)

More oral/aural assessment is desirable in English (3)

At moderation meetings the suitability of the English work programs is judged by the nature of assessment programs. (1)

Division of the Maths I course into semester units is not appropriate. (3)

Too much emphasis is placed on essays in determining ratings at foreign language moderation meetings. (1)

4. Miscellaneous

Some teachers are not dedicated to their tasks. (2)

Assignments favour students with good libraries at home. (2)

The TE Score should be used by tertiary institutions only, and not by employers. (1)

Tertiary institutions should interview potential entrants. (1)

Although teachers put considerable effort into achieving appropriate ratings for their students, comparable effort is not given to the assigning of Special Subject Assessments even though this is more important to potential tertiary entrants. (1)

Students in subjects studied mainly by high achievers, obtain lower ratings than they would have in other subjects, and consequently are disadvantaged when seeking employment. (1)

Assessment should not be the aim of education. (1)*

Ratings should be influenced more by subjective teacher judgements. (1)

Assessment should be tailored more to the individual needs of the students. (1)

The rank order of students within schools changes from Year 10 to Year 11. (1)

Holidays should be semesterised. (1)

Employers should set their own tests. (1)

4. Miscellaneous (cont)

Pre-service training of teachers in testing techniques is inadequate. (1)

The public should be educated as to the meaning of assessment measures. (1)

*Assessment should not be the aim of education. (1)**

Assessment in Years 8, 9 and 10 should be investigated. (1)

APPENDIX 5

RELATIONSHIPS BETWEEN THE RECOMMENDATIONS OF THE SELECT COMMITTEE ON EDUCATION IN QUEENSLAND, AND THE DATA COLLECTED DURING THE MAPPING OF ASSESSMENT PROCEDURES IN SCHOOLS (MAPS) STUDY

In April 1978, a Select Committee on Education in Queensland was appointed by the Queensland Parliament to inquire into, report upon and make recommendations in relation to the system of education in Queensland and the extent to which it meets societal expectations. The committee chaired by Mr Ahern presented its first interim report in November 1978. This report contained 25 recommendations. Data collected in this study is relevant to the issues contained in some of those recommendations. These have been listed in *italics*, together with comments based on the findings of this study.

10.4: *That a return to external examinations should not be considered.*

MAPS: English and foreign language teachers do not favour a return to external examinations although a small number of the latter favour common testing in some aspects of their course (e.g. grammar) to facilitate the achievement of comparability. These teachers appear to have capitalised on the syllabus freedom given to schools under the system of internal assessment and feel that this would be lost by a return to external examinations.

A large number of maths teachers favoured common testing between schools, either in the form of reference tests or external exams. These teachers expressed the feeling that there was not much room for flexibility in mathematics, and that external exams were the most effective way to achieve comparability.

10.5: *That the principle of school-based assessment should be supported.*

MAPS: With the exception of the maths teachers who wanted external examinations there was widespread support among inspectors, principals and teachers for school-based assessment. Many believed that the mechanical aspects of the present moderation system were working well, although some expressed grave doubts about the extent to which the procedures achieve comparability among schools.

10.6: *That competency-based assessment be supported in principle.*

MAPS: Although no question in the MAPS study interview referred directly to competency-based assessment, the topic frequently arose during discussions. The researchers concluded from these discussions that most principals and teachers had little knowledge of the implications of competency-based assessment. Of those who espoused the virtues of competency-based assessment most admitted that they had not read any of the report of the *Review of School-Based Assessment in Queensland Secondary Schools* and the researchers felt that they showed little understanding of its implications for schools. On the other hand, several of the interviewees who had read the report of the *Review of School-Based Assessment in Queensland Secondary Schools* had reservations about its implications. They did not see any immediate benefits to schools, and they saw potential problems for both schools and syllabus committees. They believe that a lot of work will need to be done to rate students adequately on a competency scale. Of these teachers who had considered the issues involved, the reactions ranged from "hasten slowly" to "continue with the present system, making small modifications where necessary". Some interviewees expressed the view that teachers had only recently adjusted to the present system, and that it would take a lot of effort to readjust to the new system proposed by the *Review of School-Based Assessment in Queensland Secondary Schools* report

10.8: *That the Board of Secondary School Studies should take positive action to encourage the exchange of ideas on assessement strategies and take steps to ensure that the importance of avoiding over-testing is regularly brought to the attention of school inspectors and school principals.*

MAPS: In schools which have a definite policy on assessment this usually takes the form of maximum limits on the frequency and/or length of assessments.

There already exists an acute awareness, particularly amongst principals, of the negative effects of over-testing. In general, both teachers and students are reasonably happy with the present emphasis on assessment, at least for the subjects in the study.

10.9: That the school year be re-organised in conformity with the semesterised curriculum.

MAPS: Although this topic was not covered by the interview schedules, in the MAPS study, an open-ended question on suggested changes to the present system was included. At this point, a number of teachers made spontaneous comments in favour of semesterised holidays.

10.10: That small schools receive special assistance with the administration of school-based assessment as a matter of urgency.

MAPS: Only one of the four small schools visited provided information concerning a specific problem directly attributable to its size. In another small school, however, a lot of criticism was levelled at the assessment procedures used in nearby small schools and secondary departments. Such problems were said to result from the inexperience of teachers in these smaller schools, and the unfamiliarity of most principals of secondary departments with secondary assessment procedures.

10.11: That a significant degree of flexibility in curriculum design be retained.

MAPS: It appears that teachers of English and foreign languages are using the present flexibility to advantage. Many mathematics teachers reported that the present division of semester units did not give them much flexibility. These mathematics teachers felt that more flexibility would be achieved if they could teach work when it was relevant, instead of in the strict sequence demanded by the semester system as this sequence is an artificial one rather than the one which is most suitable to development of concepts.

10.12: That a 'Curriculum Task Force' re-write each syllabus, specifying objectives in terms of cognitive skills and practical skills, as well as a quantitative expression of core content.

MAPS: Some teachers and principals commented that even if specific criteria were delineated for their subject areas, it would still be difficult to gauge acceptable standards of competence.

10.16: That the present moderation system be replaced by a system of accreditation of work programmes and certification of reported standards of achievement backed by a programme of periodic reference testing.

MAPS: Considerable concern was expressed by some respondents about the nature of reference tests. This concern centred around the quality of the tests, their content (as a limited subset of course content), who would set the tests and how the results would be used.

10.17: That the introduction of the system and any subsequent changes or modifications must be preceded by a comprehensive programme of public education, and a well planned programme of pre-service and in-service training for teachers.

MAPS: Most of the teachers who had considered the implications of the report *A Review of Schools-Based Assessment in Queensland Secondary Schools*, stated the view that adequate pre-service and in-service training was necessary before any major change is implemented. Two teachers in a sample of approximately 80 saw the public relations exercise as an attempt to indoctrinate them and the public into accepting a system which would add further complications with no added advantages. These two teachers believed that the decision to implement the recommendations of the above report had already been made without references to teacher opinion.

10.21: That in reporting levels of competency, a system of verbal as opposed to numerical or alphabetical categories be used.

MAPS: Most of the teachers who had considered the alternatives, did not see any advantage in verbal categories. While two respondents supported the use of verbal categories, many others felt that the level of misunderstanding among the public would be just as great as for numerical ratings. These teachers felt that the requirement of writing the description out fully and the large number of criteria involved would place an enormous clerical burden on school staff at a time when the clerical burden is already large.

10.23: That the TE Score system should be continued and supported.

MAPS: A surprisingly large number of teachers and subject masters appeared to have little knowledge of, or interest in, how the TE Score is calculated. Principals and teachers were asked if any particular groups (based on size, subject area, ability level, homogeneity, or any other factor) were unfairly affected by the assigning of Special Subject Assessments (SSAs) and the subsequent calculation of TE Scores. Although quite a few respondents indicated that they believed that some of the students in their subject were unfairly affected, few could give any logical reason based on the processes involved in its calculation. However some felt that ASAT did not have content validity for students in some subject areas because of its lack of emphasis on creativity andaesthetic appreciations in literature and art.

Interviewees were also asked whether it was possible to manipulate the SSA to advantage some students, and if so, how it could be done and to what extent they thought it was being done. Most interviewees believed that it was possible to manipulate the scores, and pointed out that in so doing, other students would be disadvantaged. However few respondents appeared to understand the processes sufficiently to give a procedure which would advantage a particular group of students (e.g. those students near the tertiary cut-off points). In fact, the researchers concluded that many of the suggestions made would have acted against the interest of those for whom benefit was intended.

Most respondents claimed that they did not attempt to manipulate scores themselves, but many believed that attempts at manipulation did occur in other schools. A few admitted that they did manipulate the scores of some students in their school. In two schools, some embarrassment was admitted with respect to the success of the manipulation.

APPENDIX 6

RELATIONSHIP BETWEEN THE RECOMMENDATIONS OF THE *REVIEW OF SCHOOL-BASED ASSESSMENT IN QUEENSLAND SECONDARY SCHOOLS* AND THE DATA COLLECTED IN THE *MAPPING OF ASSESSMENT PROCEDURES IN SCHOOLS* (MAPS) STUDY

On 20 February 1976, the Board of Secondary School Studies commissioned a review of school-based assessment in Queensland secondary schools. In April 1978, the committee appointed for this purpose published its report which included a number of recommendations concerning aspects of the secondary system. Data collected in this study is relevant to the issues contained in some of these recommendations. These have been listed in *italics*, together with comments based on the findings of this study.

P10: To make more efficient provisions for a curriculum organised in semester units, the school year should be re-organised into two discrete semesters, a two week break between semesters and an eight school day mid-semester break in each semester. Three days of each mid-semester break should be used for Teachers' Meetings.

MAPS: Although this issue was not specifically dealt with in the interviews, quite a few respondents spontaneously suggested altering holiday dates to fit in with the semester system.

P11: (iii) The syllabus should contain a statement of the competency criteria to be applied in each of the first three areas of general objectives to determine the student's level of competence. The specifics of these criteria should form part of the school's accredited program.

MAPS: Some respondents commented that specific competency criteria would be difficult and time consuming to establish. Of the teachers who appeared to have considered the issues involved in the preparation of syllabus containing competency-based criteria, the reactions ranged from "hasten slowly with the introduction of the new system" to "continue with the present system, making small modifications where necessary". A few interviewees expressed the view that teachers had only recently adjusted to the present system, and that it would take a lot of effort to readjust to the new system proposed by the committee.

P14: The Board should support the following statement of the purposes of assessment, viz. the purposes of assessment are:

- *to enable the student to assess his/her learning and to provide information which can be used to correct deficiencies in his/her learning;*
- *to enable the teachers and student to modify teaching/learning programs to suit the needs of the students;*
- *ultimately to provide the student and parents with information concerning the student's achievement and potential;*
- *to assist students and parents in determining future education and employment pathways; and*
- *to provide, for other educational institutions and employers, an indication of the suitability and readiness of the student to undertake further education and/or employment.*

MAPS: Principals and teachers were given a copy of these five purposes of assessment. They were then asked to consider whether:

- their testing program
- moderated school assessments (Board ratings)
- TE Scores
- school reports

facilitate, militate against or have no effect on the achievement of these purposes of assessment.

The majority of principals and teachers did not support the view often espoused that, under the present system, the testing program and exit achievement measures were militating against the achievement of the objectives of assessment. Instead, respondents generally supported the view that these four aspects of assessment were facilitating the achievement of these purposes.

P15: *The present norm-based assessment procedures in Years 10, 11 and 12 should be replaced by competency-based procedures.*

MAPS: Most interviewees admitted that they had not read either the report of the committee or the summary of its recommendations that has been distributed to schools. The researchers observed that teachers largely seemed to have little knowledge of the implications for its implementation. Of those who appeared to be reasonably familiar with its recommendations, about half liked the idea of competency-based rather than norm-based assessment. However, most considered the comparability problems will be just as great as at present and felt that in addition the workload on teachers will be greatly increased. Of those who had considered the implications of the committee's recommendations, most felt that its implementations would not alleviate the present problems associated with the production of exit measures of achievement at the Year 12 level in schools.

P16: *Assessment in each year of secondary school studies for all subjects (Board, Board-approved and school) should be made for the relevant dimensions of achievement in each subject in terms of the following competency categories:*

Highest level of competence
Highly competent
Competent
Limited competence
Very limited competence

Should the evidence of a student's competence in a given area of achievement or a given subject preclude the assigning of even the grade "very limited competence", the student should not be graded. In this event, the word "ungraded" should be recorded against that area of achievement or that subject.

MAPS: Some respondents commented that the use of verbal categories to report competence would be more meaningful to users of the measures and free schools from the need to conform to a normal distribution. However, other respondents felt that the verbal competency categories merely represented a change in labels, rather than a fundamental change in their meaning, with the disadvantage that there would no longer be checks on the distributions.

P26: *To assist in the maintenance of state-wide achievement standards and the maintenance of such standards across time, the spirit of para. 6.15 of the Radford Report should be endorsed. A policy should be adopted by using Competency Reference Tests in Board Subjects. The sole objective of these tests should be to assist schools in determining standards of performance relative to each level of competency in a subject. Such tests will be an invaluable aid to teachers in determining competency standards in smaller schools. Under no circumstances should the results of Competency Reference Tests appear on Board or School Certificates or reports. Initially Competency Reference Tests might be restricted to the Senior Secondary School with extension downwards to Year 10 depending upon the subsequent advice of the Accreditation and Certification Committee.*

MAPS: A number of respondents considered that reference tests would focus on aspects of the syllabus which were easy to assess. Thus, the flexibility allowable under present syllabuses would be lost if reference tests were used because teachers would focus only on the content to be covered and the types of assessment used by the reference test. However, some respondents desired the introduction of reference tests to guide them in determining standards.

P27 and P28

P27: *For the purpose of determining order of merit for entry to Tertiary Institutions, the Tertiary Entrance Score should be retained, though the Board should continue its research into the efficacy of the Tertiary Entrance Score as a method of ranking students.*

P28: *In calculating the Tertiary Entrance Score use should be made of ASAT, or of a comparable test, as at present. However the Board should continue its research into the use and efficacy of such a moderating instrument.*

MAPS: Concern was expressed by some respondents about the appropriateness of the use of a single measure such as the TE Score in determining tertiary entrance. A few of these, also expressed concern about the appropriateness of the Australian Scholastic Aptitude Test (ASAT) as a scaling device for some subjects.

A31: *In their assessment programs, schools should be encouraged to place more emphasis on diagnosis with remedial feedback, where the results obtained do not contribute to end of semester competency classification. The emphasis should be considered in the accreditation of school programs in each Board subject.*

MAPS: The data collected in 1978 for the present study contrasted markedly with that collected in 1974 for the Schools Under Radford Study. In 1974, teachers and administrators reported that they made little use of assessment to provide feedback to students, although they appreciated the importance of feedback in the learning process. At that time they felt that the demands of the system for reliable summative assessments prevented them from including formative assessment in their teaching programs. However, the data collected in the present study indicated that in 1978 formative assessment was much more widely used.

M23: *While choices of assessment techniques should remain largely a prerogative of the schools, the end of semester assessment in each subject should be based on information collected in a variety of ways (e.g. assignment, oral work, tests etc). Overdependence upon regular testing or oneshot examinations should be discouraged and taken into consideration in the accreditation of school programs in each subject.*

MAPS: In English and foreign languages, schools are using a wide range of assessment techniques. However, in most schools, mathematics teachers usually base their ratings on two formal tests for each semester. It was only in the computer mathematics unit of Maths I that another assessment technique (assignment) was used.

M39: *The present Senior Certificate and Tertiary Entrance Statement should be replaced with a single Senior Secondary School Certificate issued at the end of the Year 12, which should indicate ...*

(iii) The broad criteria indicating the respective levels of overall competency ...

MAPS: Some respondents indicated that they believed the development of categories indicating respective levels of overall competency would considerably increase the work load of teachers.

M45: *In its forward planning the Research Committee should be requested to give a high priority to undertaking ...*

(iii) a study of the work loads imposed upon both students and staff through continuous assessment and the implications thereof.

MAPS: The majority of principals, teachers and students commented that the present assessment load on students was generally reasonable. Some teachers indicated that they had reduced the assessment load on students in their subject because they had felt that it had been excessive in previous years.

A few teachers expressed concern that the introduction of competency-based assessment would increase their work load.

APPENDIX 7

SUGGESTED FURTHER RESEARCH

In the views of the authors, research in the following areas would add significantly to the understanding of assessment procedures in Queensland secondary schools. Where applicable, suggested approaches have been developed.

1. The Tertiary Entrance Score

The preparation of Special Subject Assessment (SSA) was discussed in Section 5 of Chapter IV. The data showed that there was considerable concern among principals and teachers that some groups of students may be unfairly affected in the calculation of TE Scores. The respondents isolated three areas of concern, the nature of the subject, the ability level of students studying the subject, and the size of the group studying the subject.

The research suggested here looks at the last two of these areas of concern. All three concerns are discussed by McGaw, Warry and McBryde (1975) and Keeves, McBryde and Bennett (1977), although the treatment used in these studies differs from that suggested here.

The first step in the calculation of the TE Score is to adjust the SSA which is the achievement score assigned by the school to each student in each Senior subject that he studied. The parameters used in the adjusting equation are:

- the particular student's SSA (S)
- the mean SSA for all students doing that subject in the school ($\bar{S}$)
- the standard deviation of all these students' SSAs (sd(S))
- the ASAT mean of all these students ($\bar{A}$)

The adjusting equation is: $A = \bar{A} + 12 \times \frac{(S - \bar{S})}{sd(S)}$ where "A" is the adjusted school assessment in subject S.

If however, a different group of students had studied a particular subject, then these parameters would be expected to have different numerical values. Thus, it follows that "A" also would have a different numerical value. One of the criteria for a valid measuring device is that the score received by a particular student is independent of who else is studying that subject, i.e. the candidature of the subject. If "A" has changed significantly, causing a resulting change in the TE Score received by a particular student, then this assumption has not been met and the situation is unfair, in that student scores are dependent on the candidature of the subject taken.

Hence, it appears desirable to investigate the effect on TE Scores of changing the candidature of a subject. Because of its very nature, this exercise can only be done artificially. It is suggested that this should be investigated by comparing what has happened with what would have happened had the candidature of the subject been different.

If the present method of calculating TE Scores is shown to be unfair, then an alternative should be developed and tested. The present system adjusts student assessments using the ASAT mean. The assumption underlying this process is that there should be a reasonable correlation between the student achievements and ASAT scores in a particular subject. Data indicates that for some subjects these correlations are very low (unpublished data collected by the Research Branch in 1970 and replicated in the A.C.T. by ACER in 1976). In some cases, student achievement in such subjects has a fairly good correlation with student achievement in subjects with a large reasonably hetrogenous candidature. If this were so, the two subjects could be considered to be conceptually linked. Thus, if a small subject were related to ASAT via a conceptually linked subject with a large hetrogenous population, then a more equitable situation may result. The concept of "piggy-backing" (i.e. linking a subject with a small candidature to a conceptually similar subject with a larger candidature) has been suggested by Keeves in personal communications, but still remains virtually unexplored. If the present system is unfair, then procedures for "piggy-backing" should be developed, and their effectiveness determined.

The steps in the proposed study are:-

1. To give a more detailed theoretical discussion of the present procedures, and of the effect of changing the candidature of a subject. In particular the effect of restricting candidature to an academically high achieving subgroup (H) will be discussed.

2. To discuss two criteria for high subjects -

    (i) comparison of State-wide ASAT means of high subjects and conceptually linked subjects; or

    (ii) comparison of the performance in the conceptually linked subjects of the subgroup studying the high subject, with the performance of the other students in the conceptually linked subject.

3. To discuss subjects reputed to be high, e.g. Maths II and foreign languages and to investigate possible link subjects, i.e. Maths I and English. Subjects that don't lend themselves to linking should be identified.

4. To analyse data relating to Maths I and Maths II. This will involve the following steps (using Maths II students as a subgroup of Maths I students) -

    (i) Test statistically the ASAT means in Maths I and Maths II;

    (ii) Test statistically the mean Maths I SSA of the Maths II subgroup and the mean Maths I SSA of the whole group;

    (These two steps are to establish eliteness)

    (iii) Calculate the Maths I "A" values of students in the subgroup (H) when they are treated as part of the whole group;

    (iv) Calculate the Maths I "A" values of students in the subgroup (H) when treated as if they were the entire group;

    (v) Investigate differences in "A" values to see if there is a consistent pattern from school to school and to determine their magnitude. (Ideally the magnitude of the differences should be expressed in terms of changes to TE Scores, and the consequent effect on careers pursued. It is recommended that this additional step be undertaken.)

    (vi) Calculate the Maths II "A" values using results obtained from the whole group's performance in Maths I;

    (vii) Calculate the Maths II "A" values using the Maths II data;

    (viii) Compare the results in (vi) and (vii). A criterion which could be used to test the relative effectiveness of the two methods would be the score of these students in Maths I. It is hypothesised that students should obtain similar adjusted scores for the high subject and the conceptually backed subject and hence the most appropriate adjustment device will be the one which gives each student a similar score in each subject studied.

2. The Validity and Reliability of School Assessment

While assessment should not be allowed to dominate the education process in schools, it does currently play an important role. Parents, tertiary institutions, and employers want schools to assess their students (i.e. summative assessment) and assessment can itself be a tool in the teaching process (i.e. diagnostic and formative assessment). Consequently, it would seem important to ensure that such assessment is of suitably high quality. But what constitutes good assessment?

Brown comments that:

> *a test must possess two characteristics if it is to be considered a good measuring instrument: the test must be reliable and it must be valid.*
>
> (Brown, p 51)

These terms are defined as follows in the Australian Item Bank, Handbook for Mathematics Item Bank (Australian Council for Educational Research):

> *Validity*
> *Although there are different types of 'validity', probably the most relevant for classroom teachers is content validity - that is, the extent to which a test item measures performance in the subject matter of the course that has been taught. To determine whether items are valid for a particular group, the educational objective being tested by each item should be examined. Teachers must make their own judgements on the relevance of items for their students.*
>
> *Reliability*
> *Like validity, the term 'reliability' has several interpretations. One is that reliability is a measure of consistency over a period of time; that is, if the same test is administered to the same students some time later, would the rank order of students be unchanged? Another interpretation is that of internal consistency - that is, are all the items in the test measuring the same broad objectives?*
>
> (Australian Item Handbook, p. 15)

Detailed descriptions of the different types of validity and reliability are given by Theobald (pp. 53-56).

However, the School Assessment Procedures Booklet No. 2 (Department of Education, Queensland, 1971, p. 49) points out that:

> *the development of challenging and valid multiple choice items is a difficult and time consuming process which requires considerable application and skill.*
>
> (School Assessment Procedures Booklet No. 2, p. 49)

The same difficulty and time factor apply to the construction of good quality items of other types as well. Many assessment experts believe that, at the present time, a large number of teachers do not possess the necessary skills to produce valid and reliable test instruments. For example, Elley and Livingstone note that:

> *internal assessment without external examination will stretch the competence of teachers in testing and observing, in marking and scaling, in judging and interpreting. While many teachers are competent to do the work required, it is clear that much more will need to be done to assist the professional growth of all teachers in these respects.*
>
> (Elley and Livingstone, p. 45)

In addition, the pressure of other professional tasks often prevents teachers from devoting more time to test construction. Therefore, it appears likely that teacher made items may not be of the high standard desired. Data collected in this study indicate that summative assessment programs used in schools rely heavily on teacher-made items. (See Table 13 in Section 2 of Chapter III). Consequently, it would seem appropriate to carry out research into the quality of summative assessments used in schools. The need for valid and reliable assessment instruments will be even greater if schools are required to provide competency-based instead of norm-based measures.

A study could be undertaken at a particular level (e.g. Year 12) for a range of subjects using a variety of assessment techniques.

The measures discussed below could be investigated.

Content validity could be examined by comparing the emphases placed on section of content in:

(i) the M1 Form;
(ii) the teaching program; and
(iii) the assessment program.

Construct validity could be investigated by validation of school assessments against recognised tests.

It is not possible to study predictive validity in a short-term study. The time which must elapse between the measurements of assessment on the predictor variable and on the criterion variable should be at least a year if reasonable data are to be obtained.

Test reliability could be investigated using split-halves and Kuder-Richardson formulae.

Assessments could be scored by several markers to establish scorer reliability.

Standard error of measurement could be calculated to establish the stability of student scores.

Item analysis could be performed to investigate:

. the difficulty of items

. the discrimination of items

. the effectiveness of distractors in multiple choice items.

3. Formative Assessment

Further research should be undertaken to examine the extent and nature of formative assessment used by teachers, including the extent to which tests and assignments comprising the summative assessment are used for formative purposes. It is envisaged that the most appropriate method of collecting data for this study would be through the use of a classroom observation technique in which the teachers would not be aware of which aspect of their teaching was being studied.

4. Study of Assessment at the Junior Level

Some concern was expressed with respect to the assessment procedures used in Years 8, 9 and 10. A pilot study would be required to determine the significant issues. It is anticipated that the issues would be different from those explored at the Year 12 level because the nature of moderation is different, and there is no pressure on students at Year 10 to meet the requirements of tertiary entrance.

5. School Reports

The researchers noted considerable diversity among schools in both what is reported, and how this information is communicated to parents and employers. A systematic study of the preferences of those who provide the reports (school administration and teachers) and the needs of those for whom the information if prepared (parents and employers) would be desirable. This study should provide schools with information which would assist them in improving the quality of their reporting.

6. Assigning of Special Subject Assessments (SSAs) by schools

The procedures used in assigning Special Subject Assessments (SSAs) to students varied considerably both between subjects and between schools.

In some cases, the score was based on a rigid application of a formula prescribed within the school by which weighted semester scores were combined and standardised. At the other extreme, the score was assigned almost entirely on the basis of subjective judgement. It is hypothesised that the shape of the distribution would depend on the method used - the empirical method producing a normal distribution, and the subjective method probably producing a range of distribution shapes. It is expected that these latter distributions would tend to be flatter (rectangular) or possibly bi-modal. Differences in the distributions will affect the values of TE Scores obtained.

It is recommended that research be undertaken into the methods used in assigning SSAs and their effects on the calculation of TE Scores.

7. Extension of the MAPS Study

Finally, it is suggested that the present study be extended to include non-state schools and other subject areas.

Policies of the Board of Secondary School Studies affect all secondary schools in Queensland. Many previous studies into the implementation and operation of internal assessment in Queensland have collected data relating to all subject areas in both state and non-state schools. However, the data collected in this study related only to three subject areas in 12 state secondary schools.

This extension, together with the present study would provide a comprehensive picture of present practices and problems relating to assessment, and would allow comparisons to be made with the data from previous studies.

APPENDIX 8

THE BOARD AND ITS CERTIFICATES

Introduction

In response to criticisms that it was not providing sufficient information to those involved in moderation procedures, the Board produced a series of information kits (Board of Secondary School Studies, 1976). This appendix reproduces much of Unit 3 of this series. The section "The Need for Certificates" has been omitted.

The Purposes of Assessment

Assessments of students' achievements in secondary school studies have a number of purposes which are deeply ingrained both in educational practice and in societal expectations of secondary education. Some of these purposes are -

- to provide, for the teacher and the student, information which can contribute to an evaluation of the teaching/learning process;
- to provide, for the student, the parent and other interested people, information which can assist in determining educational and vocational goals;
- to provide, for the student, official information concerning his/her achievements in secondary school studies which the student may regard as his personal document to use as he or she wishes.

Assessment of students' achievements is, and always has been, essentially a school responsibility. As a part of the normal process of teaching, each school plans a program of tests and assignments designed to give teachers an insight into the progress their students are making. This program will vary in form and scope according to the subject, the school and the purpose. As with other aspects of teaching no one technique can be identified as the best for all circumstances.

When the purpose of assessing the achievments of students is to assist in the teaching/learning process, the information gained is of concern only to the student, the parent and the teacher. Most benefit is derived when assessments are made at various stages of a course of study and when the information gained is fed back into the process.

When the purpose of assessing achievement is to provide, for the student, official information concerning his/her achievement in secondary school studies, the information gained will be used by the student as and when he chooses. It is in his interests that an assessment is based on a summation of the information gained throughout the course of study. Coming as it does at the end of the course of study, the information cannot be fed back into his secondary school stage, but assessments made for the purpose of certification are intended to assist the student when he leaves secondary school. A school-leaver needs such a statement of achievement; if this statement is not provided by the school system, other institutions will be forced to meet the need in a way which they will determine to meet their own needs.

The Board has no doubt that the interests of the student are best served when the information on which certificates of achievement are based, is provided by the school.

The Assessments Made

Assessments of student achievement are made at the end of the Grade 10 year, and at the end of each semester for Grades 11 and 12. The assessments made in each subject at the end of the Grade 10 year are recorded on the Junior Certificate. These assessments are reported as grades of achievement based on a descending numerical rating 7 to 1. Grades of achievement are similarly reported at the end of each semester in Grades 11 and 12. A further assessment, the Tertiary Entrance Score, is based upon performance in 20 semester units over the two year period, grades 11 and 12.

Grade 10	JUNIOR CERTIFICATE (7 to 1 Ratings)
Grades 11 - 12	
. Semester 1	
. Semester II	SENIOR CERTIFICATE (7 to 1 Ratings)
. Semester III	
. Semester IV	
Grades 11 - 12	TERTIARY ENTRANCE STATEMENT (TE SCORE)

THE ASSESSMENTS MADE

A grade of achievement is reported as a descending numerical rating 7 to 1. In this way a student's achievement in a given subject is compared with the achievements of all the other students in the State who are studying that subject at the same time. Any such grade of achievement refers to achievement in a particular subject and provides the student with information on his performance in that subject alone.

The general ability of students may differ from subject to subject. The system of grading employed takes no account of any such differences. Grades are allotted in the same proportion for each subject. Accordingly it must not be assumed that a grade in one subject can be equated to the same grade in another subject.

The percentage range of student receiving each rating on the 7 to 1 scale is as follows:

Grade	7	6	5	4	3	2	1
% range	2-6	6-12	10-20	30-50	10-20	4-12	0-6

In practice, the distribution which is aimed at is:

%	5	10	20	40	15	7	3

A Tertiary Entrance Score (TE Score) is derived from special subject assessments made at the end of the Grade 12 year, which are scaled against the Australian Scholastic Aptitude Test taken in September by all Grade 12 students.

An individual student's TE Score is based upon the aggregate of the adjusted school assessments in the student's best 20 semester units, with the provision that four semester units in each of three subjects must be included.

In 1976, TE Scores will range from 990 downwards, in intervals of 5, to 690, i.e. 990,985, 980 ... 690. A TE Score has a relative significance only. It applies to a band of students and indicates the relative position of that band of students in an order of merit list compiled from re-scaled school assessments.

Tertiary Entrance Scores are used by all tertiary institutions in Queensland as an instrument to select those students who qualify for places in them.

Moderation of Standards of Assessment

Policies relating to the issue of Board Certificates are based upon the knowledge that, at the present time, school assessments are acceptable in the community only when safeguards are seen to operate.

These safeguards apply to -

(a) the quality of the courses of study;

(b) the quality of the programs of assessment;

(c) the comparability of the ratings recorded.

The Board has a statutory responsibility to institute such safeguards. This responsibility is exercised through a system of moderation of the standards of assessment.

The Radford Committee made its position clear when it stated -

"Whatever the nature of internal examining, it must be-

(i) acceptable to teachers and parents by providing evidence that a sound education has been successfully completed;

(ii) acceptable to tertiary institutions by providing evidence of adequate preparation for and likely success in tertiary studies;

(iii) acceptable to employing authorities by providing evidence of competence to undertake particular kinds of employment."

The Committee recommended therefore that certificates be based upon school assessment "provided that there are certain safeguards."

Their proposals for instituting those safeguards included the establishment of a statutory body to be called the Board of Secondary School Studies with responsibilities covering the whole range of secondary education, in particular courses of study, the nature of certificates and moderation.

Following the Governments' adoption of the recommendations of the Radford Report, the Board was constituted as a statutory authority to implement those recommendations. Two important changes which have been implemented by the Board have been the change to the use of school assessment for the award of Senior Certificates and the change in the purpose of those certificates. Acceptance of these changes by the community has been based upon an understanding that certain safeguards would operate.

The public's interest in the operation of these safeguards is reflected in the legislation under which the Board is constituted. The Board is bound to fulfil its obligations under the Act and also to operate within the powers and functions prescribed.

Given that there is the need for a system of moderation, the Board has a responsibility to determine procedures that take into account both reasonable community expectations concerning standards of assessment and reasonable school expectations concerning responsibility for the determination of course of study and programs of assessment.

The procedures adopted for moderation of the standards of assessment are seen by the Board as providing, within the limits of the available resources, a reasonable balance between community expectations on the one hand and school expectations on the other. Since the expectations of one will at times be at variance with those of the other, it is to be expected that in the implementation of these procedures there will be some tension. A certain amount of this tension will be the normal product of growth and development. The amount of tension becomes serious, however, when there is misunderstanding of the process.

Establishing a Distribution of Ratings

The school makes an initial proposal for a distribution of ratings to be allotted to its students in each subject. In doing so, it has to make arbitrary judgements in its use of the rating scale. It will nevertheless have several sources of information on which to base this proposal. Within the school, the most important of these is the relative quality of work of individual students. Performance of its students in previous years or previous semesters is a valuable guide in that they provide a base line.

The main form of assistance provided to the school by the Board is the district moderation meeting. There are three types of meetings - one to look at programs of work and of assessment, one to look at test instruments before they are applied and one to review scripts and proposed distributions of ratings. These are used in different ways throughout the State.

Where a moderation meeting is held prior to the making of an initial proposal, the advice received from such a meeting can be used as a reliable guide.

The participants at the moderation meeting held to review the distributions proposed will express an opinion on the distribution in the light of the quality of the work submitted by the school. The consensus of the meeting is reported both to the School Moderator and the Chief Moderator. The School Moderator considers the advice and submits a subsequent proposal to the Chief Moderator.

The Chief Moderator recommends the acceptance of non-acceptance of the proposed distribution. In practice, the Chief Moderator consults with the School Moderator before recommending non-acceptance in order to provide the school with an opportunity to amend its proposal.

APPENDIX 9

CALCULATION OF THE TERTIARY ENTRANCE (TE) SCORE

In Queensland, entrance to most tertiary institutions and some faculties within those institutions is determined by the Tertiary Entrance Score (TE Score).

At some tertiary institutions various faculties require entrants to have suitably high moderated school assessments (ratings) in certain subjects in addition to a certain TE Score. (The moderated school assessment is a rating on a 7-point scale in each subject for each semester. After approval by the Board of Secondary School Studies, these appear on the Senior certificate issued at the end of Year 12.) However, for most potential tertiary entrants, the TE Score has far more effect on determining their future than do these ratings.

The procedure by which TE Scores are computed is set out below.

(a) Collection of School Assessments

For each student, schools provide an assessment, on a 1-99 scale, on each subject studied by the student, together with the number of semesters for which the student has studied the subject. These are assigned by the school using internal assessment scores and other data considered relevant using a method determined by the school. While the absolute value of these is unimportant, they are meant to reflect the relative position of students in each subject.

(b) Aptitude Measure

All students in Year 12 in State and non-State schools are required to sit for the Australian Scholastic Aptitude Test (ASAT).

(C) Adjustment of Subject Assessments

To take account of differences in ability between groups of students studying different subjects, school assessments for each subject are adjusted so that their mean is equal to the mean score on ASAT obtained by the students studying that subject. This adjustment relates the level of assessments awarded to the level of ability of the students involved.

The distribution of assessments in each subject within each school is adjusted to make the standard deviation equal to twelve. This ensures that all subjects are given the same weighting in the aggregate assessment which will finally be formed.

Where there are fewer than ten students who had studied a particular subject, it is proposed that the level of assessments should be adjusted to the mean performance on ASAT but that the standard deviation should not be altered.

(d) Aggregation of School Assessments

An aggregate of the assessments in each student's best 20 semester units is calculated, with the condition that at least three of the subjects have been studied for four semesters. For each subject, the adjusted school assessment is multiplied by the number of semesters for which that subject has been studied, thus giving a weighted school assessment for each of the subjects studied.

The weighted assessments obtained for the student's best 20 semester units of work are then summed to obtain a single assessment.

(e) Rescaling of Aggregate School Assessments

Within each school, the aggregate school assessments is then rescaled to the mean and standard deviation of the ASAT scores obtained by its students. The formula used for this rescaling is:

$$Z' = \bar{Y} + S_Y \left(\frac{X - \bar{X}}{S_X}\right)$$

where Z' is the student's rescaled school assessment

$\bar{Y}$ is the school mean on ASAT

$\bar{X}$ is the mean on the aggregate school assessments

X is the student's aggregate school assessment

and S_X and S_Y are the school's standard deviation for aggregate assessments and ASAT

An individual student's performance on ASAT will not contribute directly to his rescaled school assessment. The role of ASAT is only to remove between-school and between-subject differences in assessments.

This rescaling produces a rank order for all Year 12 students in the State who sat for ASAT.

(f) TE Scores are Assigned

The rank order list of students is divided into bands of students. All but the top band contains a number of students equal to 0.5 per cent of the 17-year-old population of the State. The top band contains 1.0 per cent. In 1978, there were approximately 38,000 17-year-olds in Queensland. Thus, 380 students lay in the top band and were assigned a TE Score of 990. Each other band contained 180 students, and were assigned TE Scores of 985, 980, 975, ... The lowest TE Score assigned in each year depends on the number of students eligible for a TE Score and the 17-year-old population of the State. This varies from year to year. The lowest TE Score assigned and the number of students in each band except the top band are tabulated below.

Year	Lowest TE Score	Band Width
1974	700	180
1975	690	180
1976	665	180
1977	655	190
1978	630	190
1979	630 (estimated)	195

GLOSSARY OF TERMS

ASAT

Australian Scholastic Aptitude Test constructed by the Australian Council for Educational Research. This test is not based on a particular syllabus, and is designed to measure aptitude for further study. The test is taken by all Year 12 students in Queensland.

BOARD OF SECONDARY SCHOOL STUDIES

A statutory authority constituted under the Education Act 1964 to 1973.

Chairman - appointed by Governor in Council
Members - nominees of Director-General of Education (7)
- nominees of each university in the State (4)
- representative of colleges of advanced education (1)
- representative of teachers colleges (1)
- teachers employed in non-government schools (4)
- nominees of teacher associations (3)
- Executive Officer of the Board (ex officio)

The functions of the Board are to exercise a general overview of the secondary curriculum, and to award certificates of secondary education based on school assessments.

M1 FORM

The Board of Secondary School Studies' M1 form outlines the proposed course of study and program of assessment in a subject. An M1 for each Board subject in each school must be submitted to the district moderator at the beginning of each semester. The course of study part of the form details *the segments of knowledge to be acquired, or the various skills or attitudes to be developed, etc.* and the time/weight of the total teaching program devoted to each of these. For each assessment element, the program of assessment part of the form details the type of assessment, length (in time or number of words), its weighting and which topics are being assessed. The textbooks and source materials to be used are listed on the back of the form.

M4 FORM

The Board of Secondary School Studies' M4 form shows the proposed distribution of ratings for each Board subject in each school for that semester. The previous ratings approved for those students in that subject are included for comparison purposes. The actual scores of the students in that semester are shown along with the points of division between one rating and the next.

MODERATION MEETING

A moderation meeting is a meeting of teachers concerned with moderation of the assessment of student achievement in a particular subject area(s). At present there are three key moderation times in the year: at the beginning (February); at the middle (July); and at the end (December).

The February meeting provides an opportunity to ensure that teachers are like-minded concerning the objectives of a course of study, the skills required, the pattern of assessment and the weighting of the various points of the course.

The July meeting provides another opportunity for reviewing courses of study and assessment programs for the semesters just beginning. In addition, the July meeting is organised to provide principals with written opinions of teachers from the schools concerning the relative standards of students' work from their own schools when compared with the rest of the district or State.

MODERATION SYSTEM

The system whereby school assessment results are reviewed in relation to the assessment results of students in other secondary schools in Queensland in a particular subject area for the awarding of junior certificates and senior certificates.

'OTHER TEACHER'

An 'other teacher' is a person other than the subject area co-ordinator who teaches a Year 12 class in the particular subject. (This term was developed for use in this study.)

RATINGS

A grade of achievement is reported as a descending numerical rating on a seven to one scale. In this way, a student's achievement in a given subject is compared with the achievements of all other students in the State who were studying that subject at the same time. Any such grade of achievement refers to achievement in a particular subject and provides the student with information on his performance in that subject alone.

SCHOOL MODERATOR

The school principal (or his/her nominee) is the person with ultimate responsibility for moderation of the assessment of student achievement in all subject areas in the school.

SPECIAL SUBJECT ASSESSMENT (SSA)

The Special Subject Assessment for a student in a subject is the achievement score assigned to the student on the basis of his performance in that subject. Schools are allowed a great deal of freedom in how they arrive at this score. The Board of Secondary School Studies each year compiles an Order of Merit list in which Year 12 students are ranked according to Special Subject Assessments that have been scaled against the ASAT test. The purpose of rescaling these assessments is to effect adjustments which render them comparable both from subject to subject within a school and from school to school within the State.

SUBJECT ADVISORY COMMITTEE

Subject Advisory Committees are appointed by the Board for subjects or groups of subjects studied or proposed for study as Board subjects in Queensland secondary schools.

The functions of a subject advisory committee are:

(a) to make recommendations to the Board concerning the broad framework of the syllabus or syllabuses in the subject or group of subjects for which the committee is appointed;

(b) to make recommendations to the Board with respect to the matter of syllabuses in subjects proposed to the Board by schools, subject associations and other interested bodies where such matter is referred to it by the Board; and

(c) to make recommendations to the Board on such other matters as the Board refers to it.

SUBJECT AREA CO-ORDINATOR (SAC)

The Subject Area Co-ordinator in a school is the person responsible for moderation of the assessment of student achievement in a particular subject area in the school. Where there is no subject master or acting subject master for that subject in the school, a Year 12 teacher generally takes on this role. (This term was developed for use in this study.)

TERTIARY ENTRANCE (TE) SCORE

With such a large number of students involved, the reporting of a complete rank order for a State Order of Merit list is not practical. Consequently, groups or bands are used for ranking. A TE Score is a centile rank based on the seventeen year old school population in Queensland in a given year. It refers to a band of students. Tertiary Entrance Scores range from 990 downwards in intervals of 5, i.e. 990, 985, 980, 975, etc., until all eligible students have been assigned a score.

REFERENCES

A review of school-based assessment in Queensland secondary schools. Board of Secondary School Studies, April 1978.

Australian item bank program - handbook for mathematics item bank. Australian Council for Educational Research, 1978.

BLOOM, B.S., HASTINGS, J.T., & MADAUS, G.F. *Handbook on formative and summative evaluation of student learning.* McGraw-Hill, New York, 1971.

BROWN, F.G. *Principles of educational and psychological testing, 2nd edition.* Holt, Rinehart & Winston, New York, 1976.

CAMPBELL, W.J., et al *Some consequences of the Radford Scheme for schools, teachers and students in Queensland.* AACRDE Report No. 7, Commonwealth Department of Education, Australian Government Publishing Service, Canberra, 1976.

CAMPBELL, W.J. & CAMPBELL, E.M. *School-based assessments: aspirations and achievements of the Radford Scheme in Queensland.* ERDC Report No. 7A, Commonwealth Department of Education, Australian Government Publishing Service, Canberra, 1978.

ELLEY, W.B. & LIVINGSTONE, I.D. *External examinations and internal assessment.* New Zealand Council for Educational Research, 1972.

FAIRBAIRN, K. & JOHNSTON, A. 'Enquiry into semesterisation', Appendix D of *A review of school-based assessment in Queensland secondary schools.* Board of Secondary School Studies, April 1978.

FAIRBAIRN, K., McBRYDE, B. & RIGBY, R. *Schools under Radford.* Queensland Department of Education and the Board of Secondary School Studies, January 1976.

First interim report of the select committee on education in Queensland - composition and functions of the Board of Secondary School Studies and secondary school assessment. Queensland Legislative Assembly, 1978.

HARTOG, P. & RHODES, E.C. *An examination of examinations.* Macmillan, London, 1936.

Information Kit, Unit 1: The Board of Secondary School Studies - an introduction. Board of Secondary School Studies, 1976.

Information Kit, Unit 3: The Board and its Certificates. Board of Secondary School Studies, 1976.

KEEVES, J.P., McBRYDE, B. & BENNETT, L.A. (ACER) 'The validity of alternative methods of scaling school assessments at the HSC level for the colleges and high schools of the Australian Capital Territory.' *Curriculum evaluation (report of the 1977 annual conference of the Australian Association for Research in Education).* Canberra, 1977.

LEWIS, D.G. *Assessment in education.* University of London Press, London, 1974.

McGAW, B., WARRY, R. & McBRYDE, B. *Validation of aptitude measures for the rescaling of school assessment.* Queensland Department of Education, 1975.

McMORROW, J.F. (Editor) *Moderation - at the crossroads?* Queensland Teachers Union, Brisbane, 1976.

Public examinations for Queensland secondary school students. Queensland Department of Education, May 1970.

School Assessment Procedures: 2. The multiple choice item. Research and Curriculum Branch, Queensland Department of Education (for the Board of Secondary School Studies), July 1971.

School Assessment Procedures: 4. Moderation within schools. Research and Curriculum Branch, Queensland Department of Education (for the Board of Secondary School Studies), May 1972.

THEOBALD, J. *Classroom testing, principles and practice.* Longman, Melbourne, 1974.